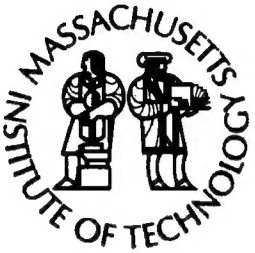
MIT/WHOI 00-23

Massachusetts Institute of Technology

Woods Hole Oceanographic Institution



Joint Program in Oceanography/ Applied Ocean Science and Engineering

DOCTORAL DISSERTATION

Genetic Diversity and Ecotypic Differentiation in the Marine Cyanobacteria *Prochlorococcus* and *Synechococcus*

by

Gabrielle Rocap

February 2000

MIT/WHOI

00-23

Genetic Diversity and Ecotypic Differentiation in the Marine Cyanobacteria *Prochlorococcus* and *Synechococcus*

by

Gabrielle Rocap

Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

and

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

February 2000

DOCTORAL DISSERTATION

Funding was provided by National Science Foundation graduate student fellowship, NASA (NAG5-3727) and NSF (OCE-9820035).

Approved for publication; distribution unlimited.

Approved for Distribution:

Laurence P. Madin

Laurence P. Madin, Chair
Department of Biology

Paola Malanotte-Rizzoli

Paola Malanotte-Rizzoli
MIT Director of Joint Program

John W. Farrington

John W. Farrington
WHOI Dean of Graduate Studies

GENETIC DIVERSITY AND ECOTYPIC DIFFERENTIATION IN THE MARINE CYANOBACTERIA *PROCHLOROCOCCUS* AND *SYNECHOCOCCUS*

by

Gabrielle Rocap

S.B. Biology
Massachusetts Institute of Technology, 1992

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

February, 2000

Signature of Author:_____
Joint Program in Biological Oceanography
Massachusetts Institute of Technology
and Woods Hole Oceanographic Institution

Certified by:_____
Sallie W. Chisholm
Professor of Civil and Environmental Engineering, and Biology
Massachusetts Institute of Technology
Thesis Advisor

Accepted by:_____
Mark Hahn
Chair, Joint Committee for Biological Oceanography
Massachusetts Institute of Technology and
Woods Hole Oceanographic Institution

GENETIC DIVERSITY AND ECOTYPIC DIFFERENTIATION IN THE MARINE CYANOBACTERIA *PROCHLOROCOCCUS* AND *SYNECHOCOCCUS*

by

Gabrielle Rocap

Submitted to the Department of Biology, Massachusetts Institute of Technology and the Woods Hole Oceanographic Institution, February, 2000 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

Primary production in the open ocean is dominated by the closely related cyanobacterial genera *Prochlorococcus* and *Synechococcus*. This thesis explores the relationship between physiological and genetic diversity in cultured isolates of *Prochlorococcus* and *Synechococcus* and examines genetic diversity in natural cyanobacterial populations to better understand the ecology of this globally important clade.

The analysis of 16S ribosomal RNA gene (rDNA) sequences of a spectrum of *Prochlorococcus* isolates from the world's oceans suggested that *Prochlorococcus* could be divided into two distinct ecotypes, designated high and low B/A because of their differing chl b/a_2 ratios. Analysis of coexisting *Prochlorococcus* populations, distinguished by their different mean chlorophyll fluorescence intensities, revealed that they were physiologically and genetically distinct. Significant differences were observed in their growth response to light and chl b/a_2 ratios and the isolates differed by 2.5% at their 16S rDNA sequences. Thus "microdiversity" commonly observed in collections of 16S rDNA sequences amplified directly from the environment, can reflect the presence of ecologically distinct populations.

To further resolve the ecotypes, the internal transcribed spacer region (ITS) between the 16 and 23S rRNAs was sequenced in isolates of *Prochlorococcus* and *Synechococcus*. ITS sequences varied widely in length and %GC content. Phylogenetic analysis identified ten sequence clusters, four of *Synechococcus* and six of *Prochlorococcus*, each of which likely corresponds to an ecologically distinct population. The high B/A ecotype contains four clusters while the low B/A ecotype contains two clusters.

Genetic diversity of natural populations of *Prochlorococcus* and *Synechococcus* was investigated in summer and winter waters of the Sargasso Sea by constructing clone libraries from PCR amplified ITS sequences, which were screened by restriction fragment length polymorphism (RFLP). Based on their phylogenetic relationships the sequences were assigned to eight clusters, seven of which had been previously identified.

Thesis Advisor: Sallie W. Chisholm
Title: Professor of Civil and Environmental Engineering, and Biology, M.I.T.

This thesis is dedicated to Celeste Rocap

"Look Mom, I did it!"

ACKNOWLEDGMENTS

My advisor, Penny Chisholm has contributed to this thesis, and to my development as a scientist, in immeasurable ways. Throughout my tenure in her lab she has constantly encouraged me to stretch my intellectual boundaries, and provided the support for me to do so. Penny's ability to focus on the big picture and her insistence on doing quality science and communicating it clearly have been an inspiration.

My committee members Colleen Cavanaugh, Jon King, Martin Polz, Mitch Sogin, and John Waterbury have been generous with their time, their laboratories, and their critical thought, all of which contributed to the final form of this thesis. I'm especially grateful to Martin, for his open office door over last two years. Thanks also go to Bill Thilly for teaching me an appreciation for the subtleties of the PCR and Dan Distel for helpful discussion in the later stages of this work.

I feel very lucky to have spent so many years in the lab with Lisa Moore, Liz Mann and Kent Cavender-Bares. By land and by sea, from the WHOI bus to the streets of Paris, they have listened to my practice talks, eaten my cooking, been my collaborators and my confidantes. I'm proud to have them as my colleagues and my friends.

Chisholm lab members past and present have been helpful on numerous occasions and generally made the lab an enjoyable place to be: Nathan Ahlgren, Lana Aref, Brian Binder, Jeff Dusenberry, Annie Rhodes, Mak Saito, Stephanie Shaw, Matt Sullivan, Claire Ting, Alex Worden, and Ena Urbach. Special thanks to Nathan who played a major role in helping me tie up many loose ends in the final months. I have also been fortunate enough to work with a number of talented undergraduates who were not only great company in the lab but taught me a lot about being a mentor: David Nadell, Leah Nichols, Nicole Herschenhous, Ethan Cohen and Welkin Pope. Ethan and Welkin both made substantial contributions to the clone libraries in Chapter 5.

Although I managed to pick a thesis topic outside the range of Sheila Frankel's considerable technical expertise, I have benefited greatly from her friendship and sage advice on all manner of life issues. Vicki Murphy cheerfully solved the increasingly bizarre funding problems I seemed to generate each semester. Carl Pabo, Michelle Mischke and my 7.014 students in the spring of 1998 helped me to be a better teacher. My fellow "general oceanographers"--Mike Atkins, Brenda Jensen, Eli Hestermann, and Nicole Poulton--always made me feel welcome in Redfield and my officemates Rachel Adams, Freddi Eisenberg and Megan Kogut helped liven up 48-320, especially during my final months of writing.

My family, both at home in Texas and my inlaws in Massachusetts were a large source of support, not to mention hot meals, over the years. Amazingly to me, although they didn't understand exactly what I was doing, they never doubted I could do it.

Finally, thanks to Paul, for all of the above, everything else, and more...

This work was funded by a National Science Foundation graduate student fellowship, NASA (NAG5-3727) and NSF (OCE-9820035).

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

CHAPTER ONE

Introduction

INTRODUCTION[1]

Primary production in the open ocean is dominated by unicellular cyanobacteria *Synechococcus* and *Prochlorococcus* (Waterbury, et al. 1979, Chisholm, et al. 1988). The two genera are abundant (often 10^5 cells ml^{-1}) and widespread in the world's oceans (Olson, et al. 1990a, Olson, et al. 1990b, Campbell and Vaulot 1993, Veldhuis and Kraay 1993). Together they have been shown to contribute 32 - 80% of the primary production in the oligotrophic oceans (Goericke and Welschmeyer 1993, Li 1995, Liu, et al. 1997, Veldhuis, et al. 1997).

Prochlorococcus and marine *Synechococcus* are closely related as determined by phylogenies constructed using several genes, including 16S ribosomal RNA (16S rDNA), a subunit of DNA-dependent RNA polymerase (*rpoC1*), and the chlorophyll *a*-binding antenna protein CP47 (*psbB*) (Palenik and Haselkorn 1992, Urbach, et al. 1992, Urbach, et al. 1998). These phylogenetic analyses demonstrate that *Prochlorococcus* and marine *Synechococcus* form a single lineage within the cyanobacteria even though the two genera use completely different light harvesting systems. *Synechococcus* has phycobilisomes typical of cyanobacteria, with open ocean *Synechococcus* distinguished from their freshwater and coastal counterparts by the presence of the phycobiliprotein phycoerythrin as their main accessory pigment (Waterbury and Rippka 1989). The major photosynthetic pigments of *Prochlorococcus* are divinyl chlorophyll *a* (8-desethyl, 8-vinyl chlorophyll *a*, or chl a_2), divinyl chlorophyll *b* (chl b_2) (Goericke and Repeta 1992).

[1] Because the chapters of this thesis were written to be journal articles which could stand alone, there is some overlap between this introduction and the introductory material in each chapter.

This complement of pigments is distinctly different from that of two other chlorophyll *b* containing oxygenic phototrophs *Prochloron* and *Prochlorothrix* (Chisholm, et al. 1992) congruent with their distant phylogenetic relationship (Urbach, et al. 1992). Because the chl-*b* containing organisms assigned to the order *Prochlorales* branch paraphyletically within the cyanobacterial radiation, the term prochlorophyte does not describe a phylogenetically coherent group (Pinevich, et al. 1997), thus *Prochlorococcus* is referred to as a cyanobacterium throughout this thesis.

Marine *Synechococcus* are larger than *Prochlorococcus*, ranging in size from 0.6-0.8 μm x 0.6-1.6 μm (Waterbury, et al. 1986) while *Prochlorococcus* are smaller (0.4-0.6 μm x 0.5-0.8 μm) (Lichtlè, et al. 1995). These size differences have implications for nutrient assimilation as the smaller *Prochlorococcus* have a higher surface-to-volume ratio and therefore will be diffusion limited at lower concentrations than the larger *Synechococcus* (Chisholm 1992).

Distribution in the Oceans

Prochlorococcus and *Synechococcus* each have a unique flow cytometric signature and are easily identified and enumerated *in situ* (Chisholm, et al. 1988, Olson, et al. 1990a). Using flow cytometry a large number of cell abundance measurements have been collected throughout the worlds oceans in the last ten years (recently reviewed in Partensky et. al. 1999). The two genera have overlapping but distinct distributions.

Geographically, *Synechococcus* has a horizontally broader distribution, as it can be found in waters with low temperatures and high nutrient concentrations where *Prochlorococcus* is absent (Partensky, et al. 1999a). However, in temperate oligotrophic waters where they coexist, *Prochlorococcus* has a greater vertical range, as it is observed in abundance much deeper in the water column than *Synechococcus*, because divinyl chl *b* allows it to more effectively harvest the low light intensities and blue wavelengths characteristic of the deep euphotic zone (Morel, et al. 1993, Moore, et al. 1995).

The two genera also exhibit differences in seasonal distributions in the oceanic regimes where they co-occur. In the Atlantic Ocean *Prochlorococcus* and *Synechococcus* have opposite seasonal cycles (Olson, et al. 1990b, DuRand, et al. 2000). Integrated concentrations of *Prochlorococcus* are maximal in the stably stratified, nutrient depleted waters in late summer and early fall when *Synechococcus* are at a minimum. Conversely integrated cell concentrations of *Synechococcus* are maximal in the deeply mixed waters in early spring when *Prochlorococcus* are at a minimum. A more pronounced version of this pattern is observed in the Gulf of Aqaba off the Red Sea where *Prochlorococcus* is not detectable in winter months when mixing is as deep as 300-600m (Lindell and Post 1995). Here, *Synechococcus* are dominant in the autumn and spring at the onset and the break up of stratified conditions, while *Prochlorococcus* is numerically dominant in the stratified low nutrient summer waters. In the permanently stratified subtropical Pacific at the HOT station, the setting is similar to summer conditions at the two above sites all year round. The water column never mixes below the euphotic zone, nitrogen remains

undetectable in the surface, and *Prochlorococcus* outnumber *Synechococcus* throughout the euphotic zone (Campbell and Vaulot 1993, Campbell, et al. 1997).

The water columns described above differ in a number of environmental parameters, including light availability and fluctuations, temperature, and nutrient and trace metal concentrations. Any or all of these factors may play a role in the observed distributions. However, the above discussion treats *Prochlorococcus* and *Synechococcus* as two homogenous groups, when in fact they are each composed of multiple physiological types which may react differently to each of the parameters mentioned. It is probably not possible to interpret the distributions of each genera without an understanding of these multiple groups.

Synechococcus

The name *Synechococcus* has been given to a supra-generic group of small (< 3 μm) unicellular coccoid cyanobacteria which lack sheaths and divide by binary fission in a single plane (Waterbury and Rippka 1989). Although similar morphologically, strains of the *Synechococcus* group are quite diverse, spanning a large range of DNA base ratios (39-71 mol % G+C). They have been provisionally divided into 6 genus level strain clusters (Cyanobacterium, Synechococcus, Cyanobium, marine A, marine B, and marine C) until further strain characterization and taxonomic analyses can be applied (Waterbury and Rippka 1989). Recent work using 16S rDNA sequences indicates that the six clusters are paraphyletic within the cyanobacterial radiation. However, strains from the

marine A, marine B and Cyanobium clusters form a monophyletic clade (Honda, et al. 1999). Open ocean *Synechococcus* isolates all belong to the marine A cluster, although marine A cluster strains have also been found in coastal waters. In contrast, marine B and C isolates are exclusively from coastal or brackish waters.

The marine A cluster *Synechococcus* encompass a physiologically diverse set of strains. Although all use phycoerythrin as their major light harvesting pigment, some also possess the chromophore phycourobilin (PUB) which can attach to phycoerythrin in place of phycoerythrobilin (PEB) (Ong and Glazer 1991). The relative amounts of PUB and PEB vary among strains (Waterbury, et al. 1986). In addition, some isolates are capable of a novel form of swimming motility (Waterbury, et al. 1985, Waterbury, et al. 1986). Marine A *Synechococcus* also vary in the ability to utilize organic nutrient sources and in their cell cycle behavior (Waterbury, et al. 1986, Binder and Chisholm 1995).

There is some evidence that the distinct physiological types of marine A *Synechococcus* differ in their distributions. Both low and high PUB type *Synechococcus* are present in both coastal and open ocean waters, while those lacking PUB are observed only in coastal waters (Olson, et al. 1988, Olson, et al. 1990a), likely because of the spectral advantage of PUB in the blue light dominated euphotic zone of the open ocean (Wood, et al. 1998). Viruses may also play a role in controlling the distribution and diversity, as the majority of *Synechococcus* in natural populations are resistant to their co-occurring phages (Waterbury and Valois 1993).

Marine A *Synechococcus* are also diverse genetically, exemplified by the range of mol percent G+C content of their genomes (55-62%) (Waterbury, et al. 1986). Using whole genome RFLP analysis four distinct lineages were identified, each as divergent from one another as from freshwater *Synechococcus* (Douglas and Carr 1988, Wood and Townsend 1990). Examination of strains from the California Current using *rpoC1* sequences identified two clusters, one containing high PUB isolates and the other consisting of low PUB isolates (Toledo and Palenik 1997). Each of these lineages was distinct from the typical laboratory model high and low PUB strains WH8103 and WH7803, suggesting that pigment content alone may not resolve the multiple groups of marine *Synechococcus* (Toledo and Palenik 1997). Recently *rpoC1* sequences were used to determine that motile isolates with a range of PUB contents are contained in a single clade (Toledo, et al. 1999). Finally, analyses using 16S rDNA sequences suggested that the high PUB strain WH8103 and the no PUB strain WH7805 are lineages taxonomically equivalent to the marine A and B divisions (Urbach, et al. 1998).

Prochlorococcus

Prochlorococcus are small, non-motile and, uniquely among organisms, contain divinyl chlorophylls *a* and *b* as their major photosynthetic pigments, although some isolates also have monovinyl chl *b* (Moore and Chisholm 1999). In addition, some isolates of *Prochlorococcus* possess the genes for phycoerythrin (*cpeA*, *cpeB*), (Hess, et al. 1996, Ting, et al. 1999), although its functional role is not yet clear. In the relatively short time

since the original description of *Prochlorococcus* rapid advances have been made in understanding their ecology, physiology and genetics, which is the subject of a recent review (Partensky, et al. 1999b).

The original open ocean divinyl chlorophyll containing isolate (strain SS120) was named *Prochlorococcus marinus* (Chisholm, et al. 1992). As additional isolates were characterized, their physiological and genetic diversity (see below) suggested that they might comprise more than one species. Until more formal taxonomic distinctions are made, a terminology parallel to *Synechococcus* has been adopted, with isolates referred to as strains of the genus *Prochlorococcus*.

Cultured isolates of *Prochlorococcus* have been divided into two genetically and physiologically distinct groups, named based on their most striking and easy to measure difference: the ratio of chl b/a_2 over a range of growth irradiances (Moore and Chisholm 1999). High B/A isolates have higher chl b/a_2 ratios, chlorophyll a_2-specific light harvesting efficiencies and spectrally-weighted average absorption coefficients, allowing them to grow at extremely low irradiances (less than 5 μmol Q m^{-2} s^{-1}) where low B/A isolates are incapable of growth. Conversely, low B/A isolates are able to grow maximally at higher light levels where high B/A isolates are inhibited (Moore and Chisholm 1999). Adaptation to light is not the only distinguishing feature of the two groups. They also differ in their sensitivity to copper toxicity, with low B/A isolates able

to grow at free cupric ion concentrations five times higher than those high B/A isolates can tolerate (Mann 2000).

The different physiologies of the two types of *Prochlorococcus* suggest that they might have different distributions in the field, with high B/A *Prochlorococcus* predominating in low-light waters in the deep euphotic zone. Low B/A isolates may dominate in high light surface waters where free cupric ion concentrations are higher, (at least in the summer in the Sargasso Sea (Moffett 1995)). Consistent with this proposed distribution, chl b/a_2 ratios in natural waters have been observed to increase with depth, and the range is larger than observed for any single isolate of *Prochlorococcus* (Goericke and Repeta 1993). Further evidence suggesting that ecologically distinct populations occupy different depths in the euphotic zone comes from copper additions to natural populations from the Sargasso Sea. In bottle incubations *Prochlorococcus* taken from deep samples disappeared immediately upon the addition of copper, while populations from surface waters remained stable, consistent with the greater sensitivity of high B/A isolates to copper (Mann 2000).

Ecotypes

Clusters of strains that are closely related genetically, as observed in *Prochlorococcus* and *Synechococcus*, are a common feature in the microbial world (Ward, et al. 1998). The multiple groups of *Prochlorococcus* and *Synechococcus* are referred to as ecotypes because it has become clear that their differing physiologies have implications for their

ecological distributions. It has been suggested that such clusters should be at the heart of a natural species concept for bacteria (Ward 1998). Without the ability to interbreed as a criterion to delineate species, microbiologists must use other means to define a "lineage evolving separately from others with its own unitary evolutionary role (ecological niche)" (Ward 1998). That is, a bacterial species is defined as an ecologically coherent population. Such populations may coexist in a guild (a group of species that exploit the same class of environmental resources) but will respond differently to environmental determinants.

Evolutionary theory predicts that ecologically distinct populations will correspond to sequence similarity clusters (Cohan 1995, Palys, et al. 1997, Ward 1998, Ward, et al. 1998). The asexual nature of bacterial evolution increases divergence between species while decreasing divergence within species. As clonal descendants of individuals with favorable mutations at functional genes sweep through a population, evolutionarily neutral mutations at all other loci are carried along as well, purging the population of diversity at all loci (Palys, et al. 1997). Thus, these periodic selection events actually increase the divergence at all loci between ecologically distinct populations. Therefore ecologically neutral mutations in rDNA or other "housekeeping genes" should be excellent markers for distinct ecotypes.

Goals of this thesis

When this research was begun, it had been established that *Prochlorococcus* were closely related to marine A *Synechococcus* and not specifically to the two other chlorophyll *b*

containing microorganisms or to chloroplasts (Palenik and Haselkorn 1992, Urbach, et al. 1992). The physiological diversity of the first few *Prochlorococcus* isolates was beginning to be explored (Morel, et al. 1993, Partensky, et al. 1993, Moore, et al. 1995). The use of ribosomal RNA as a tool to determine evolutionary relationships among cultured microorganisms was well established (Woese 1987). Studies of rDNA sequences to investigate diversity in the environment without culturing were beginning to inform microbial ecology in both terrestrial and marine environments (Giovannoni, et al. 1990, Ward, et al. 1990, Schmidt, et al. 1991, Fuhrman, et al. 1993).

The first objective of this thesis was to explore the genetic diversity of *Prochlorococcus* isolates in the context of their physiological diversity. As the major physiological distinctions between the isolates were discovered, the parallel differences in 16S rDNA sequences helped establish the two ecotypes of *Prochlorococcus* (Chapters 2 & 3). Once this correlation was established, the internal transcribed spacer region (ITS) between the 16S and the 23S rRNAs was used to further resolve the ecotypes of *Prochlorococcus* and *Synechococcus* (Chapter 4). Next, to assess how representative cultured isolates might be of natural populations, and to increase the database of sequences available for the design of ecotype specific probes, the genetic diversity at the ITS was explored in field samples (Chapter 5). Isolates and environmental sequences have been assigned to sequence similarity clusters, each of which may correspond to an ecologically distinct population. Signature sequences have been identified for many of the clusters (Chapter 6) which in the future may be used as probes in quantitative hybridizations to determine the relative

abundances of the ecotypes in natural populations. This will ultimately provide a better understanding of the distributions of these two important primary producers.

REFERENCES

Binder, B.J. and S.W. Chisholm (1995) Cell Cycle Regulation in Marine *Synechococcus* sp. Strains. *Applied and Environmental Microbiology* **61**:708-717

Campbell, L., H. Liu, H. Nolla and D. Vaulot (1997) Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991-1994 ENSO event. *Deep Sea Research* **44**:167-192

Campbell, L. and D. Vaulot (1993) Photosynthetic picoplankton community structure in the subtropical north Pacific ocean near Hawaii (station ALOHA). *Deep Sea Research* **40**:2043-2060

Chisholm, S.W. (1992) Phytoplankton Size. In: Falkowski, P. and A.D. Woodhead (ed) Primary Productivity and Biogeochemical Cycles in the Sea. Plenum Press, New York, p 231-237

Chisholm, S.W., S.L. Frankel, R. Goericke, R.J. Olson, B. Palenik, J.B. Waterbury, L. West-Johnsrud and E.R. Zettler (1992) *Prochlorococcus marinus nov. gen. nov. sp.*: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Archives of Microbiology* **157**:297-300

Chisholm, S.W., R.J. Olson, E.R. Zettler, R. Goericke, J. Waterbury and N. Welschmeyer (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**:340-343

Cohan, F.M. (1995) Does recombination constrain neutral divergence among bacterial taxa? *Evolution* **49**:164-175

Douglas, S.E. and N. Carr (1988) Examination of Genetic Relatedness of Marine *Synechococcus* spp. by Using Restriction Fragment Length Polymorphisms. *Applied and Environmental Microbiology* **54**:3071-3078

DuRand, M.D., R.J. Olson and S.W. Chisholm (2000) Phytoplankton population dynamics at the Bermuda Atlantic Time-series Station in the Sargasso Sea. *Deep Sea Research* in press

Fuhrman, J.A., K. McCallum and A.A. Davis (1993) Phylogenetic Diversity of Subsurface Marine Microbial Communities from the Atlantic and Pacific Oceans. *Applied and Environmental Microbiology* **59**:1294-1302

Giovannoni, S.J., T.B. Britschgi, C.L. Moyer and K.G. Field (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**:60-63

Goericke, R. and D.J. Repeta (1992) The pigments of *Prochlorococcus marinus*: The presence of divinyl-chlorophyll *a* and *b* in a marine prokaryote. *Limnology and Oceanography* **37**:425-433

Goericke, R. and N.A. Welschmeyer (1993) The Marine Prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. *Deep Sea Research* **40**:2283-2294

Goericke, R.E. and D.J. Repeta (1993) Chlorophylls *a* and *b* and divinyl chlorophylls *a* and *b* in the open subtropical North Atlantic Ocean. *Marine Ecology Progress Series* **101**:307-313

Hess, W.R., F. Partensky, G.W.M. van der Staay, J.M. Garcia-Fernandez, T. Börner and D. Vaulot (1996) Coexistence of phycoerythrin and a chlorophyll *a/b* antenna in a marine prokaryote. *Proc. Natl. Acad. Sci. USA* **93**:11126-11130

Honda, D., A. Yokota and J. Sugiyama (1999) Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *Journal of Molecular Evolution* **48**:723-739

Li, W.K.W. (1995) Composition of ultraphytoplankton in the central North Atlantic. *Marine Ecology Progress Series* **122**:1-8

Lichtlè, C., J.C. Thomas, A. Spilar and F. Partensky (1995) Immunological and ultrstructural characterization of the photosynthetic complexes of the prochlorophyte *Prochlorococcus* (Oxychlorobacteria). *Journal of Phycology* **31**:934-941

Lindell, D. and A.F. Post (1995) Ultraphytoplankton sucession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnology and Oceanography* **40**:1130-1141

Liu, H., H.A. Nolla and L. Campbell (1997) *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. *Aquatic Microbial Ecology* **12**:39-47

Mann, E.L. (2000) Trace Metals and the Ecology of Marine Cyanobacteria. Ph.D. Thesis. Biology, MIT and WHOI, Cambridge, MA

Moffett, J.W. (1995) Temporal and spatial variability of copper complexation by strong chelators in the Sargasso Sea. *Deep Sea Research* **42**:1273-1295

Moore, L.R. and S.W. Chisholm (1999) Photophysiology of the Marine Cyanobacterium *Prochlorococcus:* Ecotypic differences among cultured isolates. *Limnology and Oceanography* **44**:628-638

Moore, L.R., R.E. Goericke and S.W. Chisholm (1995) Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Marine Ecology Progress Series* **116**:259-275

Morel, A., Y.-H. Ahn, F. Partensky, D. Vaulot and C. Hervè (1993) *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. *Journal of Marine Research* **51**:617-649

Olson, R.J., S.W. Chisholm, E.R. Zettler and E.V. Armbrust (1988) Analysis of *Synechococcus* pigment types in the sea using single and dual beam flow cytometry. *Deep Sea Research* **35**:425-440

Olson, R.J., S.W. Chisholm, E.R. Zettler and E.V. Armbrust (1990a) Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. *Limnology and Oceanography* **35(1)**:45-58

Olson, R.J., E.R. Zettler, M. Altabet, S.W. Chisholm and J. Dusenberry (1990b) Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Research* **37**:1033-1051

Ong, L.J. and A.N. Glazer (1991) Phycoerythrins of Marine Unicellular Cyanobacteria II Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. *Journal of Biological Chemistry* **266**:9515-9527

Palenik, B. and R. Haselkorn (1992) Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. *Nature* **355**:265-267

Palys, T., L.K. Nakamura and F.M. Cohan (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *International Journal of Systematic Bacteriology* **47**:1145-1156

Partensky, F., J. Blanchot and D. Vaulot (1999a) Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. In: Charpy, L. and A.W.D. Larkum (ed) Marine Cyanobacteria. Bulletin de l'Institut Océanographique, Monaco, special issue N° 19 (632 pages), p 457-475

Partensky, F., W.R. Hess and D. Vaulot (1999b) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews* **63**:106-127

Partensky, F., N. Hoepffner, W.K.W. Li, O. Ulloa and D. Vaulot (1993) Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. *Plant Physiol.* **101**:285-296

Pinevich, A.V., S.G. Averina and N.V. Velichko (1997) Another view on the role of Photosynthetic Pigments in Taxonomy of Oxygenic-Phototrophic Bacteria: Proposed rejection of the order *Prochlorales* Florenzano, Balloni, and Materassi 1986 (Emend. Burger-Wiersma, Stal, and Mur 1989), and the Family *Prochloraceae* Florenzano, Balloni, and Materassi 1986, and the Family *Prochlorotrichaceae* Burger-Wiersma, Stal, and Mur 1989. *International Journal of Systematic Bacteriology* **47**:1264-1267

Schmidt, T.M., E.F. DeLong and N.R. Pace (1991) Analysis of a Marine Picoplankton Community by 16S rRNA Gene Cloning and Sequencing. *Journal of Bacteriology* **173**:4371-4378

Ting, C., G. Rocap, J. King and S.W. Chisholm (1999) Characterization of phycoerythrin genes in the chlorophyll a_2/b_2-containing prokaryote *Prochlorococcus* sp. MIT9303. In: Garab, G. (ed) Photosynthesis: mechanisms and effects. Kluwer Academic Publishers, The Netherlands, p 225-228

Toledo, G. and B. Palenik (1997) *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. *Applied and Environmental Microbiology* **63**:4298-4303

Toledo, G., B. Palenik and B. Brahamsha (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. *Applied and Environmental Microbiology* **65**:5247-5251

Urbach, E., D. Distel, D.J. Scanlan, J.B. Waterbury and S.W. Chisholm (1998) Rapid diversification of marine picophytoplankton with dissimilar light harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *Journal of Molecular Evolution* **46**:188-201

Urbach, E., D. Robertson and S.W. Chisholm (1992) Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature* **355**:267-270

Veldhuis, M.J.W. and G.W. Kraay (1993) Cell abundance and fluorescence of picophytoplankton in relation to growth irradiance and nitrogen availability in the Red Sea. *Neth. J. Sea. Res.* **31**:135-145

Veldhuis, M.J.W., G.W. Kraay, J.D.L. Van Bleijswijk and M.A. Baars (1997) Seasonal and spatial variation in phytoplankton biomass, productivity and growth in the northwestern Indian Ocean: the southwest and northeast monsoon, 1992-1993. *Deep Sea Research* **44**:425-449

Ward, D.M. (1998) A natural species concept for prokaryotes. *Current Opinion in Microbiology* **1**:271-277

Ward, D.M., M.J. Ferris, S.C. Nold and M.M. Bateson (1998) A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiology and Molecular Biology Reviews* **62**:1353-1370

Ward, D.M., R. Weller and M.M. Bateson (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**:63-65

Waterbury, J.B. and R. Rippka (1989) Subsection I. Order Chroococcales Wettstein 1924, Emend. Rippka et al., 1979. In: Staley, J.T., M.P. Bryant, N. Pfennig and H.J. G. (ed) Bergey's Manual of Systematic Bacteriology v. 3. Williams and Wilkins, Baltimore, p 1728-1745

Waterbury, J.B. and F.W. Valois (1993) Resistance to co-occuring phages enables marine *Synechococcus* communities to to coexist with cyanophages abundant in seawater. *Applied and Environmental Microbiology* **59**:3393-3399

Waterbury, J.B., S.W. Watson, R.R.L. Guillard and L.E. Brand (1979) Widespread occurrence of a unicellular marine planktonic cyanobacterium. *Nature* **277**:293-294

Waterbury, J.B., S.W. Watson, F.W. Valois and D.G. Franks (1986) Biological and ecological characterization of the marine unicellular cyanobacteria *Synechococcus*. *Can. Bull. Fish. Aquat. Sci.* **214**:71-120

Waterbury, J.B., J.M. Willey, D.G. Franks, F.W. Valois and S.W. Watson (1985) A cyanobacterium capable of swimming motility. *Science* **30**:74-76

Woese, C.R. (1987) Bacterial evolution. *Microbiological Reviews* **51**:221-271

Wood, A.M., D.A. Phinney and C.S. Yentsch (1998) Water column transparency and the distribution of spectrally distinct forms of phycoerythrin containing organisms. *Marine Ecology Progress Series* **162**:25-31

Wood, A.M. and D. Townsend (1990) DNA Polymorphism within the WH7803 serogroup of marine *Synechococcus* spp. (Cyanobacteria). *Journal of Phycology* **26**:576-585

CHAPTER TWO

Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* Ecotypes

Lisa R. Moore* Gabrielle Rocap* and Sallie W. Chisholm

*co first authors

1998 Moore, L. R., G. Rocap & S. W. Chisholm. Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* ecotypes. *Nature* 393 464-467

Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes

Lisa R. Moore*†, Gabrielle Rocap*†‡§ & Sallie W. Chisholm†‡

† *Department of Civil and Environmental Engineering,* ‡ *Department of Biology, 48–425 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

§ *MIT/Woods Hole Oceanographic Institute Joint Program in Oceanography, Woods Hole, Massachusetts 02543, USA*

* *These authors contributed equally to this work*

The cyanobacterium *Prochlorococcus*[1,2] is the dominant oxygenic phototroph in the tropical and subtropical regions of the world's oceans[1,3,4]. It can grow at a range of depths over which light intensities can vary by up to 4 orders of magnitude. This broad depth distribution has been hypothesized to stem from the coexistence of genetically different populations adapted for growth at high- and low-light intensities[4–6]. Here we report direct evidence supporting this hypothesis, which has been generated by isolating and analysing distinct co-occurring populations of *Prochlorococcus* at two locations in the North Atlantic. Co-isolates from the same water sample have very different light-dependent physiologies, one growing maximally at light intensities at which the other is completely photoinhibited. Despite this ecotypic differentiation, the co-isolates have 97% similarity in their 16S ribosomal RNA sequences, demonstrating that molecular microdiversity, commonly observed in microbial systems[7–12], can be due to the coexistence of closely related, physiologically distinct populations. The coexistence and distribution of multiple ecotypes permits the survival of the population as a whole over a broader range of environmental conditions than would be possible for a homogeneous population.

Using sea-going flow cytometry for studying picoplankton populations, we and others[4,13,14] have observed multiple populations of *Prochlorococcus* in single water samples, as distinguished by their chlorophyll fluorescence intensities. These populations could be derived from the mixing together of genetically identical *Prochlorococcus* cells which have acclimated to different past light

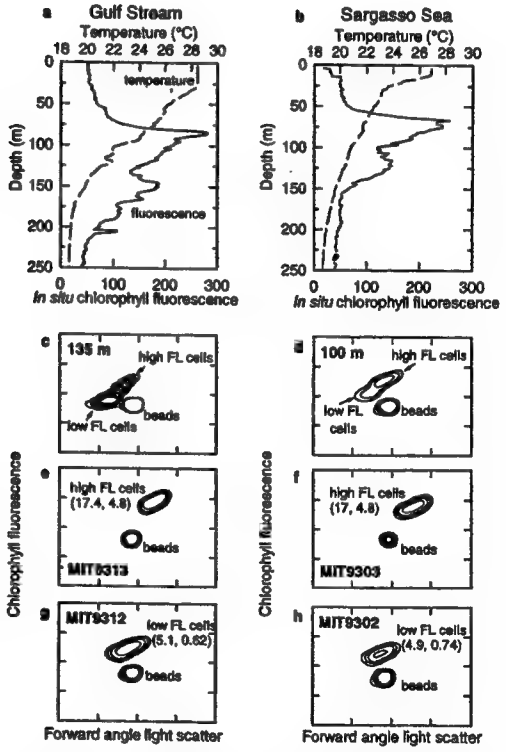


Figure 1 Properties of the euphotic zone and flow cytometric signatures of the *Prochlorococcus* populations and isolates. **a**, **b**, The physical features of the water columns were similar at the Gulf Stream station, 37° 30.8′N, 68° 14.4′W, and the Sargasso Sea station, 34° 45.5′N, 66° 11.1′W. **c**, **d**, Flow cytometry signatures of coexisting *Prochlorococcus* populations from 135 m in the Gulf Stream and 100 m in the Sargasso Sea from which the isolates were obtained. **e**–**h**, Flow cytometry signatures of the cultured isolates maintained at an irradiance of 9 μmol Q m^{-2} s^{-1}. Numbers in parentheses refer to the mean chlorophyll fluorescence per cell (FL) and FALS per cell. Differences in the absolute values of the flow cytometry parameters between natural populations and isolates result from unmatched growth conditions.

and/or nutrient regimes, or they could represent genetically distinct populations, adapted for optimal growth under different conditions but co-occurring in environments where conditions are favourable for both. The latter interpretation has been suggested by indirect evidence from field studies[4,13,14], from physiological diversity among *Prochlorococcus* isolates[6], and from studies of the molecular phylogenetic diversity of marine cyanobacteria, which have revealed a number of closely related gene sequences[7–9]. This high microdiversity has also been observed for marine bacterioplankton[7,8,10] and microorganisms in other environments[11,12], although the ecological significance has remained unclear. Differences in RNA polymerase gene sequences have been correlated with pigment content in marine *Synechococcus*[15]; however, molecular microdiversity observed in the more widely used 16S rRNA gene has not been connected directly to physiological function in any group of microorganisms from a particular habitat. The multiple *Prochlorococcus* populations observed by flow cytometry provided an excellent opportunity for studying these relationships.

In two locations of the North Atlantic Ocean at the junction of two subsurface chlorophyll fluorescence maximum layers, we used flow cytometry to sort the high- and low-fluorescence cells from coexisting *Prochlorococcus* populations and then brought them into culture (Fig. 1). After several years of growing under identical conditions, the cells sorted from a Gulf Stream water sample (isolates MIT9312 and MIT9313) and a Sargasso Sea water sample (isolates MIT9302 and MIT9303) have maintained their differences in chlorophyll fluorescence and forward-angle light scatter (FALS) per cell (Fig. 1e–h), indicating that the phenotypic differences originally observed in these coexisting populations were due to genetic differences.

We hypothesized that the high- and low-fluorescence *Prochlorococcus* are low- and high-light adapted, respectively, because differences in flow cytometry parameters are correlated with differences in photophysiology between *Prochlorococcus* SS120, isolated from the Sargasso Sea, and MED4, isolated from the Mediterranean Sea[6]. To test this hypothesis, we compared these four isolates with respect to their light-dependent growth response. The pattern of growth rate as a function of irradiance differs greatly between each coexisting pair (Fig. 2a, b). The most striking difference is that the high-fluorescence isolates, MIT9303 and MIT9313, show complete photoinhibition of growth at irradiances where the low-fluorescence isolates, MIT9302 and MIT9312, are growing at or close to maximal rates (Fig. 2a, b). Additionally, MIT9303 and MIT9313 grow better at lower light than their co-isolates (Fig. 2a, b).

To understand better the physiological underpinnings of the differential growth response of the isolates, we compared their pigment content, absorption properties and photosynthetic performance when grown at low irradiance. The high-fluorescence isolates, MIT9303 and MIT9313, contain higher concentrations of divinyl chlorophylls *a* and *b* (chl a_2 and chl b_2) relative to their low-fluorescence co-isolates MIT9302 and MIT9312 (Table 1). They also contain higher ratios of chl b_2/a_2, resulting in higher spectrally weighted average chlorophyll a_2-specific absorption coefficients, $\bar{a}^*_{chl}$

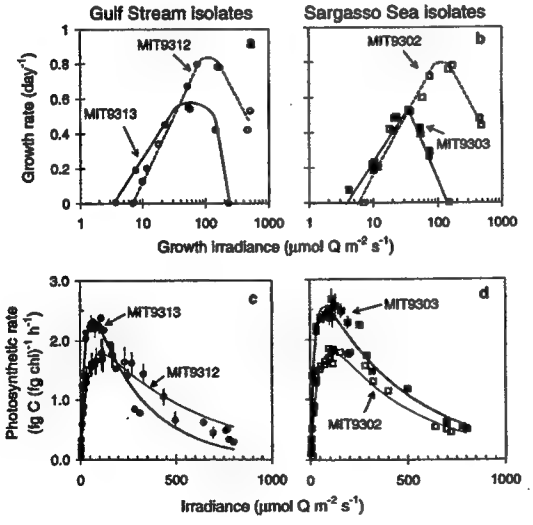
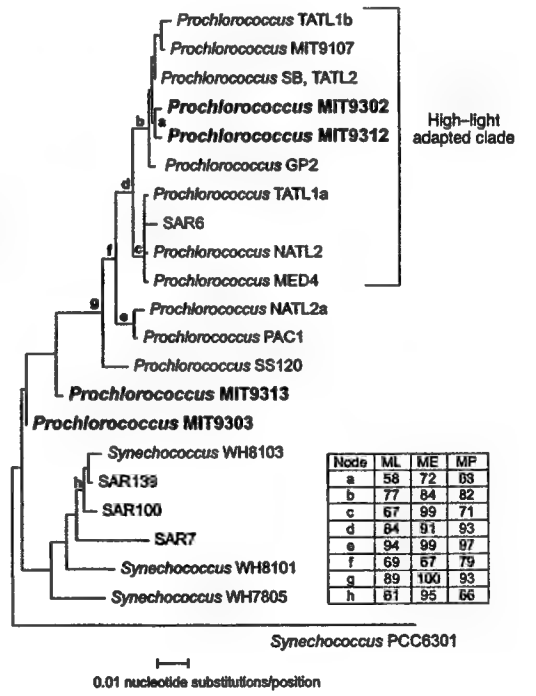


Figure 2 Growth and photosynthetic rate as a function of irradiance for the *Prochlorococcus* isolates shown in Fig. 1. **a**, **b**, Specific growth rate as a function of growth irradiance. **c**, **d**, Photosynthetic rate (normalized to chl a_2) as a function of irradiance for the four isolates when grown at an irradiance of 9 μm Q m^{-2} s^{-1}. Symbols and error bars correspond to the mean ± 1 s.e. of duplicate measurements.



Node	ML	ME	MP
a	58	72	63
b	77	84	82
c	67	99	71
d	84	91	93
e	94	99	97
f	69	67	79
g	89	100	93
h	61	95	66

Figure 3 Phylogenetic relationships of *Prochlorococcus* and *Synechococcus* isolates and environmental sequences from the Sargasso Sea ('SAR') inferred from 16S rRNA sequences. One of the three most probable trees found in a maximum likelihood analysis is presented. (Only the relative branching orders of *Prochlorococcus* isolates MED4, TATL1a and NATL2 and the environmental sequence SAR6 differ between the three trees.) Bootstrap proportions are percentages of 100 resampled data sets from analyses of maximum likelihood (ML), distance (using minimum evolution as the objective function (ME)) and maximum parsimony (MP); values below 50 are not shown. The freshwater cyanobacterium *Synechococcus* PCC 6301 was used to root the tree. Analyses with additional members of the cyanobacterial radiation resulted in essentially similar branching orders (data not shown).

Table 1 Physiological characteristics of *Prochlorococcus* isolates grown at a low growth irradiance of 9 μmol Q m^{-2} s^{-1}

	Gulf Stream		Sargasso Sea	
Parameter	MIT9312	MIT9313	MIT9302	MIT9303
Pigments				
chl a_2	3.15	5.1	3.2	4.2
chl b_2	2.08	6.0	2.1	5.0
chl b_2/a_2	0.661	1.175	0.65	1.19
Absorption				
$\bar{a}^*_{Chl}$	0.0152	0.0213	0.0147	0.0232
Photosynthetic				
α^{chla}	0.057	0.118	0.055	0.125
P^{chla}_{max}	1.8	2.4	1.8	2.54
ϕ_{max}	0.086	0.128	0.086	0.125

Chl a_2 and chl b_2 are in units of fg per cell; chl b_2/a_2 are on a weight/weight basis. $\bar{a}^*_{Chl}$ is presented in units of (m^2 (mg chl a_2)$^{-1}$). Photosynthetic parameters are in the following units: α^{chla} = fg C (fg chl a_2)$^{-1}$ h^{-1} (μmol Q m^{-2} s^{-1})$^{-1}$; P^{chla}_{max} = fg C (fg chl a_2)$^{-1}$ h^{-1}; and ϕ_{max} = mol carbon (mol Q)$^{-1}$. All values are means of duplicate cultures with coefficients of variance ≤ 10%.

(Table 1). Furthermore, the higher chl b_2/a_2 ratio of MIT9303 and MIT9313 drives a 2.1–2.3 times higher chl a_2-specific photosynthetic efficiency (α_{chla}) and results in higher maximum quantum yields (ϕ_{max}) relative to MIT9302 and MIT9312 (Fig. 2c, d; Table 1). Thus, we conclude that the isolates MIT9303 and MIT9313 are low-light adapted and MIT9302 and MIT9312 are high-light adapted.

We next explored whether the physiology of these four isolates is related to their phylogeny by sequencing their 16S rRNA genes. Although they were isolated from different waters, the two high-light adapted isolates, MIT9302 and MIT9312, have very high sequence similarity (99.7%), as do the two low-light adapted isolates, MIT9303 and MIT9313 (99.2%). The two co-isolates from each region are more disparate than the high- and low-light adapted pairs: the Sargasso Sea co-isolates MIT9302 and MIT9303 are only 97.3% similar, whereas the Gulf Stream co-isolates MIT9312 and MIT9313 are 97.7% similar. The divergence between the co-isolates is meaningful because their degrees of similarity are comparable to those between the low-light adapted *Prochlorococcus* isolates and marine isolates of phycobilisome-containing *Synechococcus* (96.9–98%). A phylogenetic tree constructed from the coexisting *Prochlorococcus* and other cyanobacterial sequences shows that the high-light adapted MIT9312 and MIT9302 cluster with the high-light adapted MED4 and two isolates from the Pacific, SB and GP2, which have low ratios of chl b/a_2 (ref. 16) (Fig. 3). The monophyly of this 'high-light adapted clade'[17] is well supported in all analyses. The low-light adapted MIT9303 and MIT9313 are on separate, more basal branches of the tree (Fig. 3). Thus, the phylogenetic relationship is correlated with the phenotypic relationship.

On the basis of the physiology and molecular phylogeny of these co-isolates, the previously characterized isolates, SS120 and MED4[6,17–19], and four additional isolates from the Pacific Ocean (ref. 20 and G.R. and S.W.C., unpublished results), it appears that at least two distinct ecotypes can be identified for isolates of the genus *Prochlorococcus*. This is analogous to two types of marine *Synechococcus*, which have different ratios of phycourobilin to phycoerythrobilin[21], and which appear to be distributed differently on the basis of their relative light harvesting abilities[22]. It has been suggested that low-light adapted, high chl b_2/a_2 *Prochlorococcus* predominate in the deeper portion of the euphotic zone where nutrients are abundant, and that high-light adapted, low chl b_2/a_2 *Prochlorococcus* predominate in the surface where nutrients are typically limiting[4–6]. This distribution of multiple *Prochlorococcus* ecotypes in the same water column would result in greater integrated production than could be achieved by a single ecotype. For example, if we estimate primary production ($P = \bar{a}^*_{chl}\phi_m C_c c_r E$) for the same number of cells (c_r) in the low light of the deep euphotic zone ($E = 10$ μmol Q m^{-2} s^{-1}) (see Methods), using the strain-specific values of $\bar{a}^*_{chla}$, ϕ_m and chl a_2 per cell (C_c), then the contribution of the low-light adapted MIT9313 to primary production would be 3.4 times higher than that of its high-light adapted coisolate MIT9312. Thus, the use of a single set of physiological parameters will result in errors in the estimates of *Prochlorococcus* primary production in the oceans.

Our results show that high phylogenetic microdiversity observed for coexisting marine cyanobacterial picoplankton[7–9] is due, in part, to physiological diversity within the *Prochlorococcus* population. Although studies using allozyme banding patterns have demonstrated ecotypic differentiation in eukaryotic phytoplankton[23,24], we have gone a step further in that we have specifically linked physiological diversity to small differences in 16S rRNA sequences. Thus, high microdiversity at the 16S rRNA locus observed in other microbial communities examined primarily by comparative sequence analysis may also reflect physiologically distinct microbial populations. The existence of multiple ecotypes in a particular environment may be a general phenomenon in the microbial world, allowing for survival over a broader range of conditions than could be achieved by a physiologically and genetically homogeneous population. □

Methods

Flow cytometry and culturing. Field populations were identified as *Prochlorococcus* on the basis of the characteristic low chlorophyll fluorescence and FALS signature on the flow cytometer[1]. Populations were designated as 'double' when bimodal signatures were visible, defined by different fluorescence and FALS signals (Fig. 1). Coexisting *Prochlorococcus* populations were sorted from each other into sterile test tubes of modified K/10-Cu media with 1.17 μM EDTA[4] using an EPICS 753 flow cytometer (Coulter Corp.). Cells were grown at 21 °C and 17 μmol quanta m^{-2} s^{-1} blue light (14:10 light:dark) immediately after isolation on the ship, and back in the laboratory at 24 °C and varying light intensities obtained using neutral density filters and cool-white fluorescent bulbs. For photosynthesis, pigment and absorption measurements, cells were grown at a low irradiance (9 μmol Q m^{-2} s^{-1}), where differences between SS120 (CCMP1375) and MED4 (CCMP1378) were apparent[6]. Cell counts for growth rate calculations and normalization of pigment data were obtained using a FACScan flow cytometer (Becton-Dickinson), and mean chlorophyll fluorescence per cell and FALS are presented relative to 0.474 μm yellow–green fluorescent beads (Polysciences). Flow cytometric data were analysed using 'CytoPC' software provided by D. Vaulot (Station Biologique, Roscoff, France).

Photosynthesis, pigment and absorption measurements. Photosynthesis–irradiance measurements were carried out on exponentially growing cells spiked with NaH^{14}CO$_3$ (0.1 μCi ml^{-1} culture; specific activity between 100,000 and 200,000 d.p.m.). The spiked culture was dispensed into glass scintillation vials (1 ml each) and incubated at 24 °C for 45 min over a range of irradiances obtained using very high output/daylight spectrum fluorescence bulbs attenuated with neutral density filters and measured with a 4π quantum light meter (model QSL-100, Biospherical Instruments, San Diego, CA). Photosynthesis was terminated and unincorporated ^{14}C removed by adding 2 M HCl (100 μl) and shaking for ~2 h under a fume hood. Samples were then analysed by scintillation spectrometry. The photosynthetic parameters, P_{max} and α, were obtained by fitting the results to the equation from ref. 25, using the curve-fitting program in SigmaPlot (Jandel Scientific), and normalized to chl a_2 and chl b_2 which were measured spectrophotometrically[26]. *In vivo* chlorophyll a_2-specific absorption, $\bar{a}^*_{chl}(\lambda)$ (m^2 (mg chl a_2)$^{-1}$), was obtained using an opal diffuser on a Beckman DU-640 spectrophotometer, and the weighted average ($\bar{a}^*_{chl}$) was calculated over the photosynthetically available radiation range (400–700 nm). Maximum quantum yield, ϕ_{max}, is calculated as the ratio of α_{chla} to $\bar{a}^*_{chl}$.

Molecular phylogenetic analysis. Genomic DNA was isolated from cultures according to standard methods[27]. 16S rDNA was amplified using the general eubacterial primers 8-27f (AGAGTTTGATCCTGGCTCAG) and 1504-1486r (CTTGTTACGACTTCACCCC). Polymerase chain reactions (PCRs) were performed in quintuplicate, using the high-fidelity polymerase *Pfu* (Stratagene). Reactions were pooled and purified using QIA quick kit (QIAGEN,

Chatsworth, CA) and cloned using the pCR-Script kit (Stratagene). A minimum of 8 clones from each culture were sequenced on an automated sequencer (LI-COR, Lincoln, NE). Sequences were aligned manually with other marine cyanobacterial sequences available in the Ribosomal Database Project[28] using the Genetic Data Environment[29]. The sequence for MIT9303 obtained previously from a PCR product[17] is contained within the sequence we report here. A total of 1,094 unambiguously aligned and determined nucleotides were used in the analyses. Phylogenetic analyses used PAUP* (version 4.0d47, provided by D. Swofford). For both distance and maximum likelihood analyses the model of nucleotide substitution used was the Hasegawa Kishino Yana 1985 model. Nucleotide frequencies and the transition transversion ratio were estimated from the data.

Received 11 November 1997; accepted 26 March 1998.

1. Chisholm, S. W. *et al.* A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334,** 340–343 (1988).
2. Chisholm, S. W. *et al. Prochlorococcus marinus* nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Arch. Microbiol.* **157,** 297–300 (1992).
3. Olson, R. J., Chisholm, S. W., Zettler, E. R., Altabet, M. A. & Dusenberry, J. A. Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Res.* **37,** 1033–1051 (1990).
4. Campbell, L. & Vaulot, D. Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (station ALOHA). *Deep Sea Res.* **40,** 2043–2060 (1993).
5. Goericke, R. & Repeta, D. J. Chlorophylls *a* and *b* and divinyl chlorophylls *a* and *b* in the open subtropical North Atlantic Ocean. *Mar. Ecol. Prog. Ser.* **101,** 307–313 (1993).
6. Moore, L. R., Goericke, R. & Chisholm, S. W. Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Mar. Ecol. Prog. Ser.* **116,** 259–275 (1995).
7. Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345,** 60–63 (1990).
8. Fuhrman, J. A., McCallum, K. & Davis, A. A. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Appl. Environ. Microbiol.* **59,** 1294–1302 (1993).
9. Palenik, B. Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Appl. Environ. Microbiol.* **60,** 3212–3219 (1994).
10. DeLong, E. F., Wu, K. Y., Prèzelin, B. B. & Jovine, R. V. M. High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371,** 695–697 (1994).
11. Ward, D., Weller, M. R. & Bateson, M. M. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345,** 63–65 (1990).
12. Amann, R., Snaidr, J., Wagner, M., Ludwig, W. & Schleifer, K.-H. In situ visualization of high genetic diversity in a natural microbial community. *J. Bacteriol.* **178,** 3496–3500 (1996).
13. Veldhuis, M. J. W. & Kraay, G. W. Cell abundance and fluorescence of picoplankton in relation to growth irradiance and nitrogen availability in the Red Sea. *Neth. J. Sea. Res.* **31,** 135–145 (1993).
14. Partensky, F., Blanchot, J., Lantoine, F., Neveaux, J. & Marie, D. Vertical structure of picophytoplankton at different trophic sites of the subtropical northeastern Atlantic Ocean. *Deep Sea Res.* **43,** 1191–1213 (1996).
15. Toledo, G. & Palenik, B. *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. *Appl. Environ. Microbiol.* **63,** 4298–4303 (1997).
16. Shimada, A., Maruyama, T. & Miyachi, S. Vertical distributions and photosynthetic action spectra of two oceanic picophytoplankters, *Prochlorococcus marinus* and *Synechococcus* sp. *Mar. Biol.* **127,** 15–23 (1996).
17. Urbach, E., Scanlan, D. J., Distel, D. L., Waterbury, J. B. & Chisholm, S. W. Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *J. Mol. Evol.* **46,** 188–201 (1998).
18. Partensky, F., Hoepffner, N., Li, W. K. W., Ulloa, O. & Vaulot, D. Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the north Atlantic and the Mediterranean Sea. *Plant Physiol.* **101,** 285–296 (1993).
19. Scanlan, D. J., Hess, W. R., Partensky, F., Newman, J. & Vaulot, D. High degree of genetic variation in *Prochlorococcus* (Prochlorophyta) revealed by RFLP analysis. *Eur. J. Phycol.* **31,** 1–9 (1996).
20. Moore, L. R. & Chisholm, S. W. Photophysiology of the marine cyanobacterium *Prochlorococcus*: ecotypic differences among cultured isolates. *Limnol. Oceanogr.* (submitted).
21. Waterbury, J. B., Watson, S. W., Valois, F. W. & Franks, D. G. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. *Can. Bull. Fish. Aquat. Sci.* **214,** 71–120 (1986).
22. Olson, R. J., Chisholm, S. W., Zettler, E. R. & Armbrust, E. V. Analysis of *Synechococcus* pigment types in the sea using single and dual beam flow cytometry. *Deep Sea Res.* **35,** 425–440 (1988).
23. Gallagher, J. C. Population genetics of *Skeletonema costatum* (Bacillariophyceae) in Narranganset Bay. *J. Phycol.* **16,** 464–474 (1980).
24. Gallagher, J. C. Physiological variation and electrophoretic banding patterns of genetically different seasonal populations of *Skeletonema costatum* (Bacillariophyceae). *J. Phycol.* **18,** 148–162 (1982).
25. Platt, T., Gallegos, C. L. & Harrison, W. G. Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. *J. Mar. Res.* **38,** 687–701 (1980).
26. Jeffrey, S. W. & Humphrey, G. F. New spectrophotometric equations for determining chlorophyll *a*, *b*, c_1 and c_2 in higher plants, algae, and natural phytoplankton. *Biochem. Physiol. Pflanz.* **168,** 191–194 (1975).
27. Ausubel, F. M. *et al. Short Protocols in Molecular Biology* (John Wiley & Sons, New York, 1992).
28. Maidak, B. L. *et al.* The RDP (Ribosomal Database Project). *Nucleic Acids Res.* **25,** 109–111 (1997).
29. Smith, S. W., Overbeek, R., Woese, C. R. & Gilbert, W. The genetic data environment: an expandable GUI for multiple sequence analysis. *CABIOS* **10,** 671–675 (1994).

Acknowledgements. We thank M. Sogin for access to an automated DNA sequencer and helpful discussion; A. Shimada for the use of unpublished sequences of *Prochlorococcus* GP2 and SB; R. Olson and E. Zettler for cell sorting at sea; H. Hsu for technical assistance; and C. Cavanaugh, H. Sosik and S. Sathyendranath for helpful consultation. This work was supported by the US National Science Foundation and the US National Aeronautics and Space Administration.

Correspondence and requests for materials should be addressed to S.W.C. (e-mail: chisholm@mit.edu).
Sequences have been deposited in GenBank under accession numbers AF053396, MIT9302; AF053397, MIT9303; AF053398, MIT9312; and AF053399, MIT9313.

CHAPTER THREE

Molecular Phylogeny of *Prochlorococcus* Ecotypes

Gabrielle Rocap, Lisa R. Moore and Sallie W. Chisholm

Reprinted with permission of the Musée Océanographique de Monaco

1999 G. Rocap, L. R. Moore, and S. W. Chisholm. Molecular phylogeny of *Prochlorococcus* ecotypes. In Marine Cyanobacteria, Charpy, L. & Larkum, A. W. D. (eds). Bulletin del'Institut Oceanographique, Monaco, special issue N°19, pp. 107-116.

Prochlorococcus is the dominant phototroph in the temperate oceans. It is found throughout the euphotic zone, at light intensities spanning over four orders of magnitude. This growth range is larger than observed for any single culture of _Prochlorococcus_ and is likely due to the co-existence of multiple ecotypes adapted for optimal growth at different light intensities. Here we report the 16S rDNA sequences of four _Prochlorococcus_ isolates from the Pacific Ocean, MIT9201, MIT9202, MIT9211 and MIT9215, which have been physiologically characterized as high or low B/A ecotypes based on their chl b/a_2 ratios and light-dependent growth responses. The three low B/A isolates are closely related and cluster together with other low B/A isolates in a previously identified "high-light adapted" clade. The fourth Pacific isolate, a high B/A ecotype, is not a member of this clade. Sequences specific to each physiological type will be useful in probing field populations to determine the spatial and temporal distribution of the ecotypes of _Prochlorococcus_.

The marine cyanobacterium *Prochlorococcus* (Chisholm et al. 1988, Chisholm et al. 1992) is abundant (often 10^5 cells ml^{-1}) and widespread in the tropical and subtropical regions of the open ocean. *Prochlorococcus* is found throughout the euphotic zone and often comprises 40-60% of the total chlorophyll in the Pacific (Chavez et al. 1991, Letelier et al. 1993, Campbell et al. 1994, Shimada et al. 1995), the north Atlantic (Goericke & Welschmeyer 1993, Veldhuis et al. 1993), and the Red Sea and Somalia Upwelling region (Veldhuis & Kraay 1993, Lindell & Post 1995).

We refer to *Prochlorococcus* as a cyanobacterium, despite its original description as a "prochlorophyte" (Chisholm et al. 1988) — referring to oxygenic photosynthetic prokaryotes that possess chl *b* and lack phycobilins (Lewin 1981). Molecular phylogenies constructed using both the 16S ribosomal RNA gene (16S rDNA) (Urbach et al. 1992) and the gene for the large subunit of RNA polymerase (*rpoC1*) (Palenik & Haselkorn 1992) demonstrate that *Prochlorococcus* is most closely related to marine *Synechococcus,* and phylogenetically distinct from other prochlorophyte lineages. The close relationship between *Prochlorococcus* and *Synechococcus* is also apparent in phylogenies constructed using the *psbA* and *psbB* genes and portions of the *petB* and *petD* genes (Hess et al. 1995, Urbach et al. 1998), and in the possession of phycoerythrin genes by some strains of *Prochlorococcus* (Hess et al. 1996, Ting et al. 1998). Further, *Prochlorococcus* has a suite of pigments distinctly different from that of two other previously designated prochlorophytes *Prochloron* and *Prochlorothrix* (Chisholm et al. 1992). It is now clear that the chl *b* phenotype has evolved independently at least four times (La Roche et al. 1996) and that these organisms, along with traditional cyanobacteria, should probably be reclassified as Oxyphotobacteria (Pinevich et al. 1997). Although phylogenetic analyses using the gene for the large subunit of Rubisco (*rbcL*) place one isolate of *Prochlorococcus* with the gamma proteobacteria (Shimada et al. 1995), this likely reflects a horizontal gene transfer event, and not the true organismal phylogeny (Delwiche & Palmer 1996). Two other *Prochlorococcus* isolates possess *rbcL* sequences which branch within the cyanobacterial clade (Pichard et al. 1997) and one

strain of *Synechococcus* has a proteobacterial-like *rbcL* sequence as well (Watson & Tabita 1996).

Using flow cytometry to study picoplankton in the field, we and others have observed distinct co-occurring populations of *Prochlorococcus*, based on their different mean chlorophyll fluorescence intensities (Campbell & Vaulot 1993, Veldhuis & Kraay 1993, McManus & Dawson 1994, Partensky et al. 1996). Although these populations could have been derived from the mixing together of genetically identical *Prochlorococcus* cells acclimated to different past light and/or nutrient regimes, several lines of evidence suggested they were physiologically and genetically distinct populations. The latter interpretation was supported by physiological experiments defining two different types of *Prochlorococcus* (Partensky et al. 1993, Moore et al. 1995). A comparative study of two isolates, the type strain SS120 (CCMP1375) and MED4 (CCMP1378), demonstrated that saturating growth irradiances are significantly lower for SS120 than for MED4: SS120 is photoinhibited above 140 $\mu E\ m^{-2}\ s^{-1}$, and ratios of chlorophyll *b* to divinyl chlorophyll *a* (chl a_2 (Goericke & Repeta 1993)) are almost an order of magnitude higher in SS120 over all irradiances measured (Moore et al. 1995). Consistent with these data, biochemical characterization of the two strains indicates that the pigment complexes of SS120 are bigger and more abundant at low light, and bind seven times more chl *b* than those of MED4 (Partensky et al. 1997).

The coexistence of physiologically and genetically distinct types of *Prochlorococcus* was definitively shown recently when we flow-cytometrically sorted and cultured the high and low fluorescence cells from coexisting *Prochlorococcus* populations in two locations of the North Atlantic Ocean (Moore et al. 1998). At low irradiances the high fluorescence isolates, MIT9303 and MIT9313, have relatively high growth rates and a high chl b/a_2 ratio, similar to SS120, whereas the low fluorescence isolates, MIT9302 and MIT9312, have a lower growth rate and a low chl b/a_2 ratio, similar to MED4 (Moore et al. 1998).

The physiology of these six isolates and four isolates from the Pacific Ocean (MIT9201, MIT9202, MIT9211, MIT9215) suggests the genus *Prochlorococcus* can be divided into at least two ecotypes based on pigment content and light-dependent physiological response (Moore & Chisholm 1999). Pigment content has also been used to distinguish isolates of *Synechococcus* (e.g. high and low PUB types; see Waterbury et al. 1986). The distinction between the two groups of *Prochlorococcus* is best defined using their chl b/a_2 ratios (Fig. 1) since at any particular irradiance there is no overlap in this parameter. One group of *Prochlorococcus* isolates, the low B/A ecotype (Moore & Chisholm 1999), cluster with chl b/a_2 ratios below 0.7 while the other four isolates, the high B/A ecotype, have higher chl b/a_2 ratios, but exhibit more variability between them (Fig. 1). High B/A isolates are also characterized by their higher chlorophyll a_2-specific light harvesting efficiencies and spectrally-weighted average absorption coefficients. High B/A isolates also show decreases in growth rate due to high light exposure at light levels where the

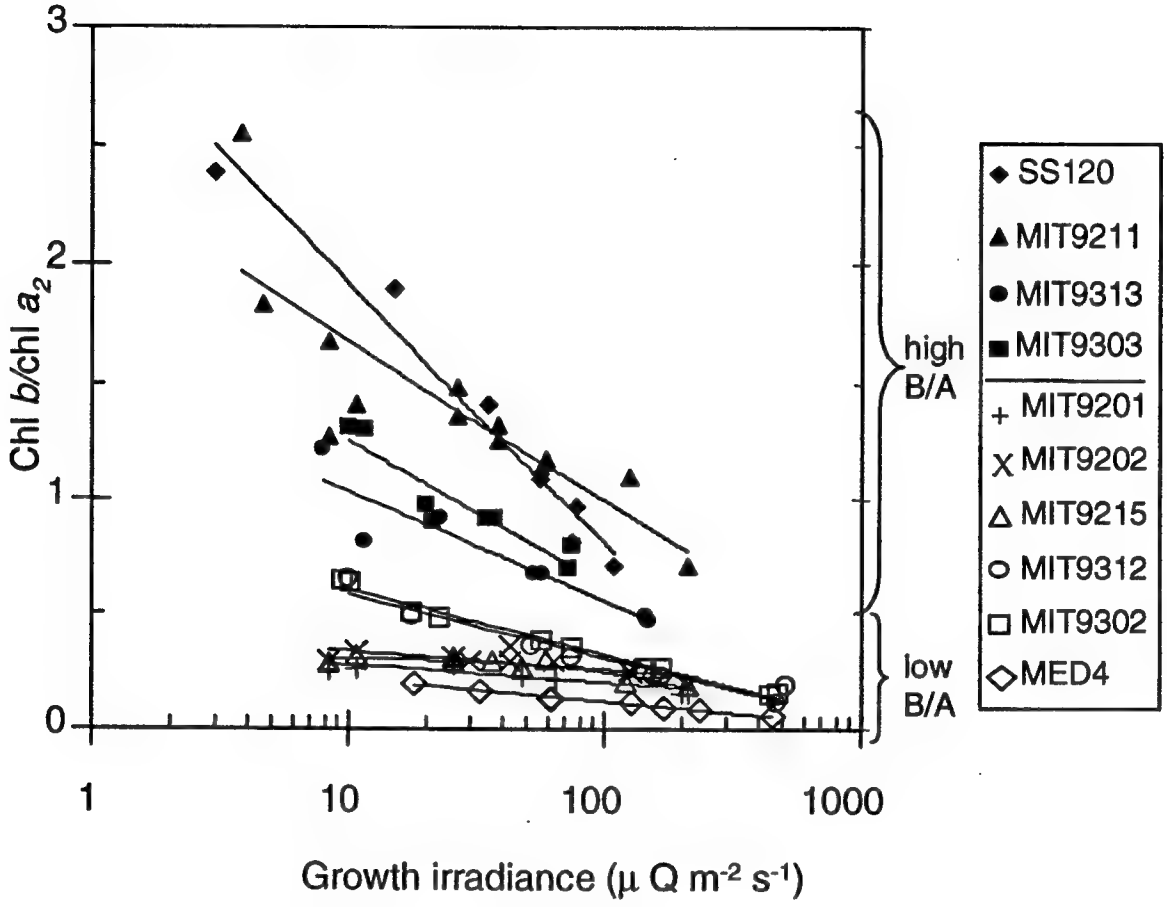


Figure 1. Chlorophyll b/a_2 ratio of different *Prochlorococcus* isolates grown over a range of light intensities (for isolation information see Moore & Chisholm, 1999). The isolates can be grouped loosely into two clusters, those with low chl b/a_2 ratios (low B/A ecotype) and those with high b/a_2 ratios (high B/A ecotype). The high B/A ecotype appears to have two sub-groups within it. Data from Moore & Chisholm, 1999.

low B/A isolates grow maximally. Further, a secondary division can be made within the high B/A isolates: when grown under high light MIT9303 and MIT9313 have higher cell-specific light harvesting efficiencies and maximum photosynthetic rates than SS120 and MIT9211 (Moore & Chisholm 1999).

Genetic analysis of *Prochlorococcus* isolates indicates that there is a correlation between photophysiology and phylogeny. Characterization of ten isolates by RFLP mapping (using probes to photosynthetic genes) groups them into two clusters, one containing the high B/A isolate SS120 and the other containing the low B/A isolate MED4 (Scanlan et al. 1996). Analysis of 16S rDNA sequences from these isolates (Urbach et al. 1998) reveals a well supported group, deemed the "high-light adapted clade" because it includes the low B/A MED4 and two Pacific isolates, SB and GP2, which have low chl b/a_2 ratios (Shimada et al. 1996). Support for this clade is further strengthened by analyses demonstrating that the low B/A isolates MIT9302 and MIT9312 (each from a coexisting pair of isolates) are also members of this clade (Moore et al. 1998).

The 16S rDNA sequences of four *Prochlorococcus* isolates from the Pacific Ocean further support the correlation of photophysiology with phylogeny. On the basis of their photophysiology three of these isolates (MIT9201, MIT9202 and MIT9215) belong to the low B/A ecotype, while the fourth (MIT9211) is characterized as high B/A (Moore & Chisholm 1999). The three low B/A isolates are more than 99% similar to each other in their 16S rDNA sequence and cluster within the "high-light adapted clade" (Fig. 2). One

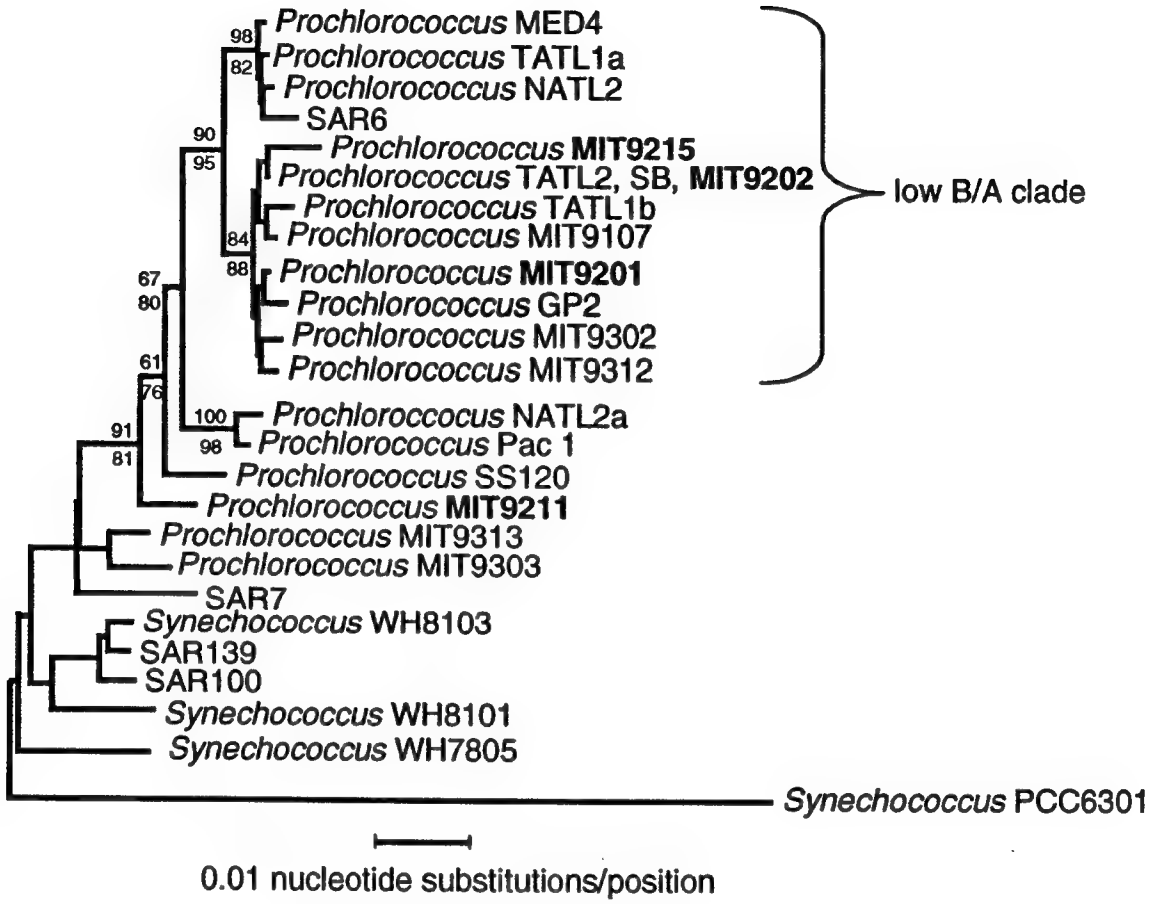


Figure 2. Phylogenetic relationships of *Prochlorococcus* and *Synechococcus* isolates and environmental sequences (SAR) from the Sargasso Sea inferred from 16S rRNA gene sequences. Sequences of Pacific isolates MIT9201, MIT9202, MIT9211 and MIT9215 (in bold) were determined using previously described methods . All analyses (tree construction and similarity values cited in text) employed 1070 unambiguously determined and aligned bases of the 16S rDNA. The phylogenetic framework was inferred by a distance analysis, using minimum evolution as the objective function. Bootstrap values from distance and parsimony analyses (above and below the branches, respectively) are presented out of 100 resampled datasets; values below 50 are not shown. The freshwater cyanobacterium *Synechococcus* PCC6301 was used to root the tree.

isolate, MIT9202, has a sequence identical to those of two other isolates, TATL2, from the Tropical Atlantic, and SB from the West Pacific, corroborating previous results demonstrating that *Prochlorococcus* from different oceanic regimes can be genetically similar (Urbach et al. 1998). Although all of the isolates which cluster in the low B/A clade are closely related (at least 98.8% similar to one another), there is a well supported subdivision within this clade. One of these subgroups contains MED4, which has the lowest chl b/a_2 ratio of all the isolates we have characterized (Fig.1), although no further physiological distinction between the two subgroups is apparent.

Congruent with the larger degree of physiological variability among isolates of the high B/A ecotype, they also exhibit greater genetic variability than the low B/A isolates. The 16S rDNA sequence similarities of MIT9211 with SS120, MIT9313 and MIT9303 are 98.6%, 98.2% and 97.8% respectively. These similarities are lower than those between the low B/A isolates, and are only slightly greater than the similarities between MIT9211 and *Synechococcus* WH8103 (97.4%). The high B/A isolates MIT9303 and MIT9313 share even higher sequence similarities to *Synechococcus* (Moore et al. 1998). The high B/A isolates do not cluster together in a single clade but are on separate, basal branches of the tree (Fig. 2), consistent with the possible subdivisions within this ecotype suggested by their photophysiology.

Future directions include further consideration of the taxonomic status of the ecotypes we have described. It is possible that the high and low B/A ecotypes represent different

species of *Prochlorococcus*. However, a definitive determination will require either sequence data from a more variable locus to provide more resolution between the high B/A isolates and marine *Synechococcus*, or axenic cultures of each ecotype in order to perform other taxonomic analyses such as DNA-DNA hybridizations and determinations of % GC content and fatty acid composition (Wayne et al. 1987, Priest & Austin 1993, Stackebrandt & Goebel 1994, Tamaoka 1994). At present we adopt the convention used for marine *Synechococcus* (Waterbury et al. 1986) and designate our isolates simply as strains of the genus *Prochlorococcus*. Although we have definitively described two ecotypes, we consider this a minimum number. As described above, there are undoubtedly subdivisions within both the high and low B/A ecotypes. Future investigations will likely reveal additional ecotypes not currently represented in culture and/or further ecotypic differentiation based on susceptibility to cyanophages, nutrient utilization efficiencies, or trace metal toxicity (Mann et al. 1997).

A second direction we are actively pursuing is determining the distributions of the two ecotypes in the field. It has been suggested that low-light adapted, high chl b/a_2 *Prochlorococcus* predominate in the deeper portion of the euphotic zone and that high-light adapted, low chl b/a_2 *Prochlorococcus* predominate in surface waters (Campbell & Vaulot 1993, Goericke & Repeta 1993, Moore et al. 1995). Analyses of natural populations, by examining sequences from environmental clone libraries (Ferris & Palenik 1998) and using molecular probes designed to take advantage of the correlation of physiology with phylogeny (N. West & D. Scanlan, pers. comm.), are consistent with

this hypothesized distribution. Repeated observations of multiple phylotypes in single water samples (Palenik 1994, Ferris & Palenik 1998, Urbach & Chisholm 1998) indicate that further work is needed to determine how the ecotypes are distributed under different environmental conditions. It is clear that the distribution of multiple *Prochlorococcus* ecotypes in the same water column would result in greater integrated production than could be achieved by a single ecotype (Moore et al. 1998). The existence of multiple ecotypes allows for survival over a broader range of conditions than could be achieved by a physiologically and genetically homogenous population and undoubtedly contributes to the dominance of *Prochlorococcus* in the world's oceans.

REFERENCES

Campbell L, Nolla HA, Vaulot D (1994) The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. Limnol Oceanogr 39:954-961

Campbell L, Vaulot D (1993) Photosynthetic picoplankton community structure in the subtropical north Pacific ocean near Hawaii (station ALOHA). Deep Sea Res 40:2043-2060

Chavez FP, Buck KR, Coale K, Martin JH, DiTullio GR, Welschmeyer NA, Jacobson AC, Barber RT (1991) Growth rates, grazing, sinking and iron limitation of equatorial Pacific phytoplankton. Limnol Oceanogr 36:1816-1833

Chisholm SW, Frankel SL, Goericke R, Olson RJ, Palenik B, Waterbury JB, West-Johnsrud L, Zettler ER (1992) *Prochlorococcus marinus nov. gen. nov. sp.*: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. Arch Microbiol 157:297-300

Chisholm SW, Olson RJ, Zettler ER, Goericke R, Waterbury J, Welschmeyer N (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. Nature 334:340-343

Delwiche CF, Palmer JD (1996) Rampant horizontal gene transfer and duplication of rubisco genes in eubacteria and plastids. Mol Biol Evol 13:873-882

Ferris MJ, Palenik B (1998) Niche adaptation in ocean cyanobacteria. Nature 396:226-228

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. Nature 345:60-63

Goericke R, Welschmeyer NA (1993) The Marine Prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. Deep Sea Res 40:2283-2294

Goericke RE, Repeta DJ (1993) Chlorophylls *a* and *b* and divinyl chlorophylls *a* and *b* in the open subtropical North Atlantic Ocean. Mar Ecol Prog Ser 101:307-313

Hess W, Weihe A, Loiseaux-De Goër S, Partensky F, Vaulot D (1995) Characterization of the single *psbA* gene of *Prochlorococcus marinus* CCMP 1375 (Prochlorophyta). Plant Mol Biol 27:1189-1196

Hess WR, Partensky F, van der Staay GWM, Garcia-Fernandez JM, Börner T, Vaulot D (1996) Coexistence of phycoerythrin and a chlorophyll *a/b* antenna in a marine prokaryote. Proc Natl Acad Sci USA 93:11126-11130

La Roche J, van der Staay GWM, Partensky F, Ducret A, Aebersold R, Li R, Golden SS, Hiller R, Wrench PM, Larkum AWD, Green BR (1996) Independent evolution of the prochlorophyte and green plant chlorophyll a/b light harvesting proteins. Proc Natl Acad Sci USA 93:15244-15248

Letelier RM, Bidigare RR, Hebel DV, Ondrusek M, Winn CD, Karl DM (1993) Temporal variability of phytoplankton community structure based on pigment analysis. Limnol Oceanogr 38:1420-1437

Lewin RA (1981) The Prochlorophytes. In: Starr MP, Stolp H, Trüper HG, Balows A, Schlegel HG (ed) The Prokaryotes. Springer, Berlin, p 257-266

Lindell D, Post AF (1995) Ultraphytoplankton sucession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. Limnol Oceanogr 40:1130-1141

Mann EL, Lam P, Hsu H, Chisholm SW, Moffett JW (1997) Does copper influence the distribution of marine cyanobacteria in the Sargasso Sea? ASLO Aquatic Sciences Meeting Santa Fe, NM

McManus GB, Dawson R (1994) Phytoplankton pigments in the deep chlorophyll maximum of the Caribbean Sea and the western tropical Atlantic Ocean. Mar Ecol Prog Ser 113:199-206

Moore LR, Chisholm SW (1999) Photophysiology of the Marine Cyanobacterium *Prochlorococcus:* Ecotypic differences among cultured isolates. Limnol Oceanogr in press

Moore LR, Goericke RE, Chisholm SW (1995) The comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. Mar Ecol Prog Ser 116:259-275

Moore LR, Rocap G, Chisholm SW (1998) Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* Ecotypes. Nature 393:464-467

Palenik B (1994) Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. Appl Environ Microbiol 60:3212-3219

Palenik B, Haselkorn R (1992) Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. Nature 355:265-267

Partensky F, Blanchot J, Lantoine F, Neveux J, Marie D (1996) Vertical structure of picophytoplankton at different trophic site of the subtropical northeastern Atlantic. Deep Sea Res 43:1191-1213

Partensky F, Hoepffner N, Li WKW, Ulloa O, Vaulot D (1993) Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. Plant Physiol 101:285-296

Partensky F, LaRoche J, Wyman K, Falkowski PG (1997) The divinyl-chlorophyll *a/b*-protein complexes of two strains of the oxyphototrophic marine prokaryote *Prochlorococcus* -- characterization and response to changes in growth irradiances. Photosynthesis Res 51:209-222

Pichard SL, Campbell L, Paul JH (1997) Diversity of the Ribulose Bisphosphate Carboxylase/Oxygenase Form I gene (*rbcL*) in natural phytoplankton communities. Appl Environ Microbiol 63:3600-3606

Pinevich AV, Averina SG, Velichko NV (1997) Another view on the role of Photosynthetic Pigments in Taxonomy of Oxygenic-Phototrophic Bacteria: Proposed rejection of the order *Prochlorales* Florenzano, Balloni, and Materassi 1986 (Emend. Burger-Wiersma, Stal, and Mur 1989), and the Family *Prochloraceae* Florenzano, Balloni, and Materassi 1986, and the Family *Prochlorotrichaceae* Burger-Wiersma, Stal, and Mur 1989. Intl J Syst Bact 47:1264-1267

Priest F, Austin B (1993) Modern Bacterial Taxonomy. Chapman and Hall, London

Scanlan D, Hess W, Partensky F, Newman J, Vaulot D (1996) High degree of genetic variation in *Prochlorococcus* (Prochlorophyta) revealed by RFLP analysis. Eur J Phycol 31:1-9

Shimada A, Kanai S, Maruyama T (1995) Partial sequence of Ribulose-1,5-Bisphosphate Caboxylase/Oxygenase and the physiology of *Prochloron* and *Prochlorococcus* (Prochlorales). J Molec Evol 40:671-677

Shimada A, Maruyama T, Miyachi S (1996) Vertical distributions and photosynthetic action spectra of two oceanic picophytoplankters, *Prochlorococcus* and *Synechococcus* sp. Mar Biol 127:15-23

Shimada A, Nishijima M, Maruyama T (1995) Seasonal appearance of *Prochlorococcus* in Suruga Bay, Japan in 1992-1993. J of Oceanography 51:289-300

Stackebrandt E, Goebel BM (1994) Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. Intl J Syst Bact 44:846-849

Tamaoka J (1994) Determination of DNA Base Composition. In: Goodfellow M, O'Donnell AG (ed) Chemical Methods in Prokaryotic Systematics. Wiley, New York, p 463-470

Ting C, Rocap G, King J, Chisholm SW (1998) Characterization of phycoerythrin genes in the chlorophyll a_2/b_2-containing prokaryote *Prochlorococcus* sp. MIT9303. 11th International Photosynthesis Congress Budapest, Hungary

Urbach E, Chisholm SW (1998) Genetic diversity in *Prochlorococcus* populations flow cytometrically sorted from the Sargasso Sea and the Gulf Stream. Limnol Oceanogr 43:1615-1630

Urbach E, Distel D, Scanlan DJ, Waterbury JB, Chisholm SW (1998) Rapid diversification of marine picophytoplankton with dissimilar light harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). J Molec Evol 46:188-201

Urbach E, Robertson D, Chisholm SW (1992) Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. Nature 355:267-270

Veldhuis MJW, Kraay GW (1993) Cell abundance and fluorescence of picophytoplankton in relation to growth irradiance and nitrogen availability in the Red Sea. Neth J Sea Res 31:135-145

Veldhuis MJW, Kraay GW, Gieskes WWC (1993) Growth and fluorescence characteristics of ultraplankton on a north-south transect in the eastern North Atlantic. Deep Sea Res II 40:609-626

Waterbury JB, Watson SW, Valois FW, Franks DG (1986) Biological and ecological characterization of the marine unicellular cyanobacteria *Synechococcus*. Can Bull Fish Aquat Sci 214:71-120

Watson GMF, Tabita FR (1996) Regulation, unique gene organization, and unusual primary structure of of carbon fixation genes from a marine phycoerythrin-containing cyanobacterium. Plant Mol Biol 32:1103-1116

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Trüper HG (1987) Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Intl J Syst Bact 37:463-464

CHAPTER FOUR

Resolution of *Prochlorococcus* and *Synechococcus* ecotypes based on pigment analysis and 16S/23S rDNA Internal Transcribed Spacer (ITS) region sequences

ABSTRACT

Cultured isolates of the marine picoplankton *Prochlorococcus* and *Synechococcus* vary widely in their pigment compositions and growth responses to light and nutrients. In order to better define the genetic variation that accompanies this physiological diversity, sequences for the 16S 23S rDNA internal transcribed spacer region (ITS) were determined in 30 *Prochlorococcus* isolates and 14 marine A *Synechococcus* isolates from around the globe. All strains examined yielded a single ITS sequence which contained two tRNA genes. Dramatic variations in the length and %GC content of the spacer were observed among the strains, particularly among *Prochlorococcus*. The 44 strains can be grouped into ten sequence similarity clusters, each of which likely represents an ecologically distinct population. The ITS enabled the identification of four clusters within marine A *Synechococcus* strains, and six clusters of *Prochlorococcus*, two within the low B/A ecotype and four within the high B/A ecotype. A better understanding of physiological and phylogenetic diversity in cultured isolates provides an important framework for studies which draw ecological conclusions about natural populations based solely on DNA sequence data.

INTRODUCTION

In open ocean ecosystems, carbon fixation is dominated by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Together they have been shown to contribute between 32 - 80% of the primary production in the oligotrophic oceans (Goericke and Welschmeyer 1993, Li 1995, Liu, et al. 1997, Veldhuis, et al. 1997). *Prochlorococcus* is

closely related to marine A *Synechococcus* based on analyses using gene sequences from 16S ribosomal RNA (16S rDNA) and a subunit of DNA-dependent RNA polymerase (*rpoC1*) (Palenik 1992, Urbach, et al. 1992). However, *Prochlorococcus* contains divinyl chlorophyll *a* (8-desethyl, 8-vinyl chlorophyll *a*, or "chl a_2") and divinyl chlorophyll *b* (chl b_2) as its major photosynthetic pigments rather than the phycobilisomes that are typical of cyanobacteria (Chisholm, et al. 1988, Chisholm, et al. 1992).

Cultured isolates of *Prochlorococcus* have been divided into two genetically and physiologically distinct groups, referred to as ecotypes because their differing physiologies have implications for their ecological distributions (Moore, et al. 1998, Moore and Chisholm 1999, Rocap, et al. 1999). High B/A isolates have larger ratios of chl b/a_2 and are able to grow at extremely low irradiances (less than 5 μmol Q m^{-2} s^{-1}) where low B/A isolates are incapable of growth. Low B/A isolates have lower chl b/a_2 ratios and are able to grow maximally at higher light intensities, where high B/A isolates are inhibited (Moore and Chisholm 1999). The ecotypes also differ in their sensitivity to copper toxicity, with low B/A isolates able to grow at free cupric ion concentrations five times higher than those high B/A isolates can tolerate (Mann 2000).

The16S rDNA sequences of the *Prochlorococcus* ecotypes correlate with their physiology. Strains of the low B/A ecotype are phylogenetically very closely related (99% similarity in 16S rDNA sequence) and cluster in a tight clade (Moore, et al. 1998, Urbach, et al. 1998, Rocap, et al. 1999). Strains of the high B/A ecotype are less similar

in their16S rDNA sequence (97-98%) and do not form a monophyletic group (Rocap, et al. 1999). High B/A *Prochlorococcus* also have a higher degree of sequence similarity to marine *Synechococcus* than the low B/A ecotype. In fact, branching orders between some high B/A *Prochlorococcus* isolates and *Synechococcus* are not well resolved using 16S rDNA sequences (Moore, et al. 1998, Urbach, et al. 1998, Rocap, et al. 1999).

Open ocean *Synechococcus* consist of a genetically and physiologically diverse set of strains. All contain phycoerythrin as their major light-harvesting pigment, and some possess the chromophore phycourobilin (PUB) which can attach to phycoerythrin in place of phycoerythrobilin (PEB) (Ong and Glazer 1991). The relative amounts of PUB and PEB vary among strains. In addition, some isolates are capable of a novel form of swimming motility (Waterbury, et al. 1985, Waterbury, et al. 1986). *Synechococcus* also vary in the mol percent G+C content of their genomes, and in the ability to utilize organic nutrient sources such as urea and cyclic AMP (Waterbury, et al. 1986).

Genetic diversity in *Synechococcus* has been examined in a few strains using16S rDNA sequences (Urbach, et al. 1998) and more extensively using *rpoC1* sequences (Toledo and Palenik 1997, Toledo, et al. 1999). Using *rpoC1* a collection of strains from the California Current could be divided into two lineages consistent with their high or low PUB amounts. However, each of these lineages was distinct from the typical laboratory model high and low PUB strains WH8103 and WH7803, suggesting that pigment content alone may not resolve the multiple ecotypes of marine *Synechococcus.* Motility,

however, does appear to be correlated with phylogeny, as all motile isolates characterized to date are closely related (Toledo, et al. 1999).

Clusters of strains that are closely related genetically, as observed in *Prochlorococcus* and *Synechococcus*, are a common feature in the microbial world (Ward, et al. 1998). It has been suggested that such sequence similarity clusters should be at the heart of a natural species concept for bacteria (Ward 1998), and that ecologically neutral mutations in rDNA or other "housekeeping genes" are excellent markers for distinct ecotypes (Palys, et al. 1997). In practice, however, as in the case of high B/A *Prochlorococcus* and *Synechococcus*, 16S rDNA can be too conserved to adequately resolve closely related yet ecologically distinct strains.

In most eubacteria the genes for ribosomal RNA are organized in operons with the genes encoding the 5S, 16S and 23S rRNAs separated by internal transcribed spacer (ITS) regions. The spacer regions probably have a role in holding secondary structure of preRNA for processing to mature rRNAs (Apirion and Miczak 1993). The spacer between the 16S and 23S rRNA genes can encode 0, 1, or 2 tRNA genes and exhibit a great deal of length and sequence variation. The ITS has been used in many bacterial groups to delineate closely related strains (Barry, et al. 1991, Gürtler and Stanisich 1996, Leblond-Bourget, et al. 1996, Christensen, et al. 1999). Because there is only one rRNA operon in *Prochlorococcus,* the ITS, like the 16S rDNA, is present in a single copy (Urbach, et al. 1998, West and Scanlan 1999).

Here we report on use of the ITS as phylogenetic tool to identify strain clusters which may represent ecologically distinct populations of marine cyanobacteria. Although isolates in culture collections may not represent the full extent of diversity due to biases introduced by isolation protocols, they provide crucial physiological information to attach to the strain clusters. By examining a wide range of physiologically diverse isolates of *Prochlorococcus* and *Synechococcus* we hope to lay the groundwork for informed studies of genetic diversity and distributions in field populations of marine cyanobacteria.

MATERIALS AND METHODS

Isolation of strains. Thirty isolates of *Prochlorococcus* from diverse oceanic regimes were employed in this study (Table 1). Isolation conditions, physiology and genetic data have been reported previously for many of the strains (see Table 1). The majority were isolated by filtering seawater through two stacked 0.6 μm-pore-size filters and enriching with nutrients (Chisholm, et al. 1992). Five were isolated by sorting on a flow cytometer (Moore, et al. 1998). Seven of the strains (SS120, SS35, SS51, SS2, MED4, SB, and GP2) have been rendered clonal by serial dilution and one (MED4Ax) has been rendered free of heterotrophic bacteria by plating (M. Saito and J. Waterbury, unpublished data).

The 30 strains of *Prochlorococcus* include the 10 isolates characterized by Moore and Chisholm (1999), the 15 isolates described in this work and 5 additional isolates. These 5 additional isolates are clonal derivatives of previously characterized strains. The original

Table 1. *Prochlorococcus* isolates used in this study.

ISOLATES	OTHER NAMES	COORDINATES	DEPTH	DATE ISOLATED	ISOLATION METHOD	ORIGINAL REFERENCE
Sargasso Sea						
SARG	LG, SSW5	29°N; 64.4°W	120 m	May '88	filtered enrichment	(Chisholm, et al. 1992)
SS120	CCMP1375	"	"	"	clonal by dilution from SARG	(Chisholm, et al. 1992)
SS35	CCMP1428	"	"	"	"	"
SS51	CCMP1376	"	"	"	"	"
SS2	CCMP1427	"	"	"	"	"
MIT9301		34.2°N; 66.3°W	90 m	Jul '93	filtered enrichment	this work
MIT9302		34.8°N; 66.2°W	100 m	Jul '93	sorted on FCM	(Moore, et al. 1998)
MIT9303		"	"	"	sorted on FCM	(Moore, et al. 1998)
MIT9401		35.5°N; 70.4°W	Surface	May '94	filtered enrichment	this work
Gulf Stream						
MIT9311		37.5°N; 68.2°W	135 m	Jul '93	sorted on FCM	this work
MIT9312		"	"	"	sorted on FCM	(Moore, et al. 1998)
MIT9313		"	"	"	sorted on FCM	(Moore, et al. 1998)
MIT9314		37.5°N; 68.2°W	180 m	"	filtered enrichment	this work
North Atlantic						
NATL1A		37.4°N; 40°W	30 m	Apr. '90	filtered enrichment	(Partensky, et al. 1993)
NATL2A		39°N; 49.3°W	10 m	"	filtered enrichment	(Scanlan, et al. 1996) (Urbach, et al. 1998)
Mediterranean Sea						
MED	DV1	43.2°N; 6.9°E	5 m	Jan '89	filtered enrichment	
MED4	CCMP1378	"	"	"	clonal by dilution from MED	(Moore, et al. 1995)
MED4Ax		"	"	"	axenic, cloned from MED4 by plating	Saito et al in prep
South Pacific						
MIT9107		15°S; 135°W	25 m	Aug '91	filtered enrichment	(Urbach, et al. 1998)
MIT9116		"	"	"	filtered enrichment	this work
MIT9123		"	"	"	filtered enrichment	this work
MIT9201		12°S; 145.4°W	Surface	Sep '92	filtered enrichment	(Moore and Chisholm 1999)
MIT9202		"	79 m	"	filtered enrichment	(Moore and Chisholm 1999)
Equatorial Pacific						
MIT9211		0°; 140°W	~83 m	Apr '92	filtered enrichment	(Moore and Chisholm 1999)
MIT9215		0°; 140°W	Surface	Oct '92	filtered enrichment	(Moore and Chisholm 1999)
MIT9321		1°N; 92°W	50 m	Nov '93	filtered enrichment	this work
MIT9322		0.3°N; 93°W	Surface	Nov '93	filtered enrichment	this work
MIT9515		5.7°S; 107.1°W	15 m	Jun '95	filtered enrichment	this work
Western Pacific						
GP2		8°N; 136°E	150m	Sep. '92	filtered enrichment, clonal by dilution	(Shimada, et al. 1995a)
SB		35°N; 138.3°E	40m	Oct. '92	filtered enrichment, clonal by dilution	(Shimada, et al. 1995b)
Arabian Sea						
AS9601		19°S; 67°E	50 m	Nov '95	filtered enrichment	(Shalapyonok, et al. 1998)

isolate of *Prochlorococcus*, SARG, which is the parent strain of the high B/A isolate SS120 (the type strain of *Prochlorococcus marinus*), has also given rise to an isolate with genetic and physiological properties which place it in the low B/A ecotype (Rippka, et al. 2000). To investigate the potential heterogeneity of the SARG culture, SARG and 3 of its other clonal derivatives were selected for sequencing. Because it will be the subject of many future investigations, *Prochlorococcus* strain MED4Ax, rendered axenic by picking a single colony from an agar plate (Saito and Waterbury, unpublished data), was also sequenced in order to compare it with its parent strain MED4.

The isolation and culture history of isolates NATL1A and NATL2A is convoluted enough to warrant a special mention. Isolates NATL1 and NATL2 were isolated from the North Atlantic in 1990 and kindly provided to the MIT *Prochlorococcus* culture collection by Fred Partensky shortly thereafter. NATL1 was originally reported to have a chl b/a_2 ratio and light dependent growth responses which placed it in the high B/A ecotype (Morel, et al. 1993, Partensky, et al. 1993). However, the isolates maintained in the Roscoff Culture Collection of Marine Phytoplankton (France) changed over time to have a phenotype similar to the low B/A ecotype (described in Rippka, et al. 2000). The altered phenotypic properties are consistent with RFLP and 16S rDNA analyses demonstrating that both Roscoff isolates are closely related to the low B/A isolate MED4 (Scanlan, et al. 1996, Urbach, et al. 1998, Rippka, et al. 2000). In contrast the isolates maintained at MIT continued to maintain a phenotype characteristic of the high B/A ecotype (L. Moore unpublished observations). The MIT NATL2 isolate yielded a 16S

rDNA sequence which branched outside the low B/A clade and was 2.5% different than the sequence from the Roscoff NATL2 (Urbach, et al. 1998). The MIT culture was designated NATL2A (Urbach, et al. 1998) and in parallel, we refer to the NATL1 culture maintained at MIT as NATL1A to distinguish it from the isolate maintained at Roscoff.

Fifteen of sixteen *Synechococcus* isolates used in this study (Table 2) are clonal and have been described previously (Waterbury, et al. 1986). Strain WH9908 was isolated from Woods Hole in April 1999 (by M. Sullivan), when the water temperature was less than 10° C and rendered clonal by picking a single colony from an agar plate. Of the 16 isolates, 14 are phycoerythrin-containing marine A cluster strains. One freshwater strain assigned to the Cyanobium cluster (PCC6307) and one strain from the marine B cluster (WH8101) are included as outgroups (Table 2).

Culture conditions. For physiology experiments, 20 ml batch cultures of *Prochlorococcus* were grown in acid washed 50 ml test tubes at 24°C on 14:10 light:dark cycle under 18 μmol Q m^{-2} s^{-1} white light. This light level is roughly equivalent to 1% of surface irradiance, which corresponds to a depth of ~100m (assuming typical oligotrophic water values for surface irradiance of I_o = 2000 μmol Q m^{-2} s^{-1} and an extinction coefficient of k = 0.045 m^{-1} (Kirk 1994)). This growth irradiance was chosen based on physiological data from ten other *Prochlorococcus* isolates (Moore and Chisholm 1999), because it was a level at which all were capable of growth and one where ecotypic differences in the chl b/a_2 ratio were pronounced.

Table 2. *Synechococcus* isolates used in this study

STRAIN	ISOLATION LOCATION	DATE	PE ?	A495/ A545	GENOME %GC	MOTILE ?
Cyanobium cluster						
PCC6307	Wisconsin Lake	1949	-	NA	69.7	-
Marine B cluster						
WH8101	Woods Hole	1981	-	NA	63.9	-
Marine A cluster						
WH6501	8.7°N, 50.8°W	1965	+	0.43	62.3	-
WH7803	33.7°N, 67.5°W	July '78	+	0.39	61.3	-
WH7805	33.7°N, 67.5°W	June '78	+	no PUB	59.7	-
WH8002	19.7°N, 92.4°W	Apr. '80	+	0.48	60	-
WH8008	19.7°N, 92.4°W	Apr. '80	+	no PUB	55.8	-
WH8012	34°N, 65°W	Jul. '80	+	0.4	62.4	-
WH8016	Woods Hole	June '80	+	0.4	55.5	-
WH8017	Woods Hole	June '80	+	0.42	54.5	-
WH8018	Woods Hole	June '80	+	no PUB	57.6	-
WH8103	28.5°N, 67.4°W	Mar. '81	+	2.4	58.9	+
WH8109	39.5°N, 70.5°W	Jun. '81	+	0.89	?	-
WH8112	36°N, 66°W	Oct. '81	+	variable	59.8	+
WH8113	36°N, 66°W	Oct. '81	+	variable	60.5	+
WH9908	Woods Hole	Apr. '99	+	?	?	-

Data in table adapted from Waterbury et al. (1986). PE indicates the presence or absence of phycoeryrthrin. A495/A545 ratio is a measure of relative amounts of pigments phycourobilin (PUB) to phycoerythrobilin (PEB).

Media was made from 0.2 μm filtered, autoclaved Sargasso Sea water enriched with "Pro2" nutrients (final concentrations 10μM NaH_2PO_4, 50μM NH_4Cl, 100μM urea, 1.17 μM EDTA, 8nM Zn, 5 nM Co, 90 nM Mn, 3 nM Mo, 10 nM Se, 10 nM Ni, and 1.17 μM Fe) (Moore and Chisholm 1999).

For DNA extraction *Prochlorococcus* were grown in 60 ml acid washed polycarbonate bottles using the media and culture condition described above. *Synechococcus* were

grown for DNA extraction in 100ml acid washed flasks in SN media under constant light (Waterbury, et al. 1986).

Growth rate, pigment and flow cytometric measurements. The physiology of fifteen isolates of *Prochlorococcus* was examined by measuring their growth rate, pigment content and flow cytometric properties. All physiology experiments were performed in triplicate (although the chl b/a_2 ratio and total chlorophyll values presented for MIT9321 represent the mean of duplicate cultures). Cells were acclimated to the experimental conditions for at least ten generations before measurements were taken. Growth rate was determined by measuring the fluorescence of each culture at the same time of day in a fluorometer (Turner model 10-AU).

For pigment measurements a known volume (18-22 ml) of exponential phase culture was filtered onto a 25 mm Whatman GF/F under low vacuum. Filters were stored in liquid nitrogen until extraction. Pigments were extracted according to established protocols (Moore, 1997, adapted from Goericke and Repeta, 1993), and quantified on a spectrophotometer (Becton Dickinson DU640). Unlike HPLC, spectrophotometric methods cannot resolve divinyl chlorophyll b_2 from "normal" monovinyl chlorophyll b, thus total ($b_1 + b_2$) values are reported. Pigment concentrations were calculated according to the trichromatic equations of Jeffrey and Humphreys (1975).

For flow cytometry, 1 ml of exponential phase culture was preserved in glutaraldehyde (final concentration 0.125%) and stored in liquid nitrogen until cell numbers and relative red fluorescence were determined by running samples on modified FACScan flow cytometer (Dusenberry and Frankel 1994). Fluorescence per cell is reported relative to 0.474 μm diameter yellow-green, polystyrene microspheres (Polysciences, Inc. Warrington, PA). Data were analyzed using Cyclops software (Cicero).

DNA isolation, PCR and sequencing. DNA was extracted from 50 ml of late exponential phase cultures using a modified CTAB/phenol chloroform protocol (Ausubel, et al. 1992). The ITS/23S fragment was amplified using primers 16S-1247f (CGTACTACAATGCTACGG) and 23S-1608r (CYACCTGTGTCGGTTT). Primer 16S-1247f was designed using availiable16S rDNA sequences from *Prochlorococcus* and *Synechococcus* (Rocap, et al. 1999) and is a perfect match only to cyanobacterial 16S rDNA sequences as judged by using the CHECK PROBE function of the latest release of the Ribosomal Database Project (Maidak, et al. 1999). Reactions were done in 25 μl volume with final concentrations of reactants as follows: 0.25 mM dNTPs, 0.1 mM each primer, 0.1-1 μg template DNA and 0.1-0.5 U of the high fidelity polymerase Pfu. Cycling parameters were 94° for 4 min followed by 30 cycles of (94° 1 min, 52°1 min, 72° 6 min) and a final extension at 72° for 10 min using either a robocycler (Stratagene, La Jolla CA) or a PTC100 (MJ Research). Amplified fragments were visualized on agarose gels. Only one band was observed from each culture. Control reactions lacking

template DNA were always performed in parallel and gave no products, as judged by inspection on agarose gels.

For sequencing, PCR reactions were performed in quintuplicate, pooled, and primers removed using Strataprep columns (Stratagene). Products were sequenced on ABI377 or ABI310 (PE BIOsystems) automated sequencers using Big Dye terminator sequencing kits according to the manufacturer's instructions. The ITS was sequenced bidirectionally using primer 16S-1247F, and primers internal to the PCR fragment: ITS-alaf (TWTAGCTCAGTTGGTAGAG), ITS-alar (CTCTACCAACTGAGCTAWA) and 23S-241r (TTCGCTCGCCRCTACT).

Phylogenetic analysis. Sequences were edited and aligned manually using the Genetic Data Environment (Smith, et al. 1994). Phylogenetic analyses and calculation of fractional similarities and % GC contents used PAUP* version 4b2a (Swofford 1999). Phylogenetic analyses employed either 103 or 146 positions of the 16S-tRNA^{Ile} spacer.

Distance trees were inferred using minimum evolution as the objective criterion and paralinear (logdet) or HKY85 distances. Distance and maximum-parsimony bootstrap analyses (1000 resamplings) were performed with heuristic searches utilizing random addition and tree-bisection reconnection (TBR) branch-swapping methods. Maximum likelihood analyses used the HKY85 model of nucleotide substitution with rate heterogeneity and empirical nucleotide frequencies. The gamma shape parameter and the

transition/transversion ratio were initially estimated from a distance topology and refined by iterative likelihood searches. Likelihood bootstrap analyses (100 resamplings) were performed with heuristic searches and tree-bisection reconnection (TBR) branch-swapping methods starting from neighbor joining tree. Phylogenetic trees were visualized with TREEVIEW (Page 1996).

RESULTS

Physiology of *Prochlorococcus* strains. Fifteen isolates of *Prochlorococcus* were grown at 18 μmol Q m^{-2} s^{-1} (roughly equivalent to the 1 % light level) and their chl b/a_2 ratio was measured. Growth rates of the isolates ranged between 0.3 - 0.5 day^{-1} (Fig. 1A). Chl b/a_2 ratios for thirteen of the isolates ranged between 0.2 and 0.5 while two, NATL1A and NATL2A, had much larger ratios (0.86 and 0.97 respectively) (Fig. 2B). In a previous study six low B/A isolates had chl b/a_2 ratios of 0.2 - 0.5 at comparable irradiance levels, while the four isolates of the high B/A ecotype had chl b/a_2 ratios which ranged from 0.9 - 1.5 (Moore and Chisholm 1999). Thus, these results suggest that NATL1A and NATL2A belong to the high B/A ecotype, while the other thirteen isolates characterized here are of the low B/A ecotype.

Total chlorophyll per cell observed in these isolates was well within the range seen previously (1-9 fg $cell^{-1}$) (Fig. 1C), as was the flow cytometrically derived chlorophyll (red) fluorescence (Fig. 1D) (Moore and Chisholm 1999). Differences in flow cytometrically derived fluorescence have been observed in natural populations (Campbell

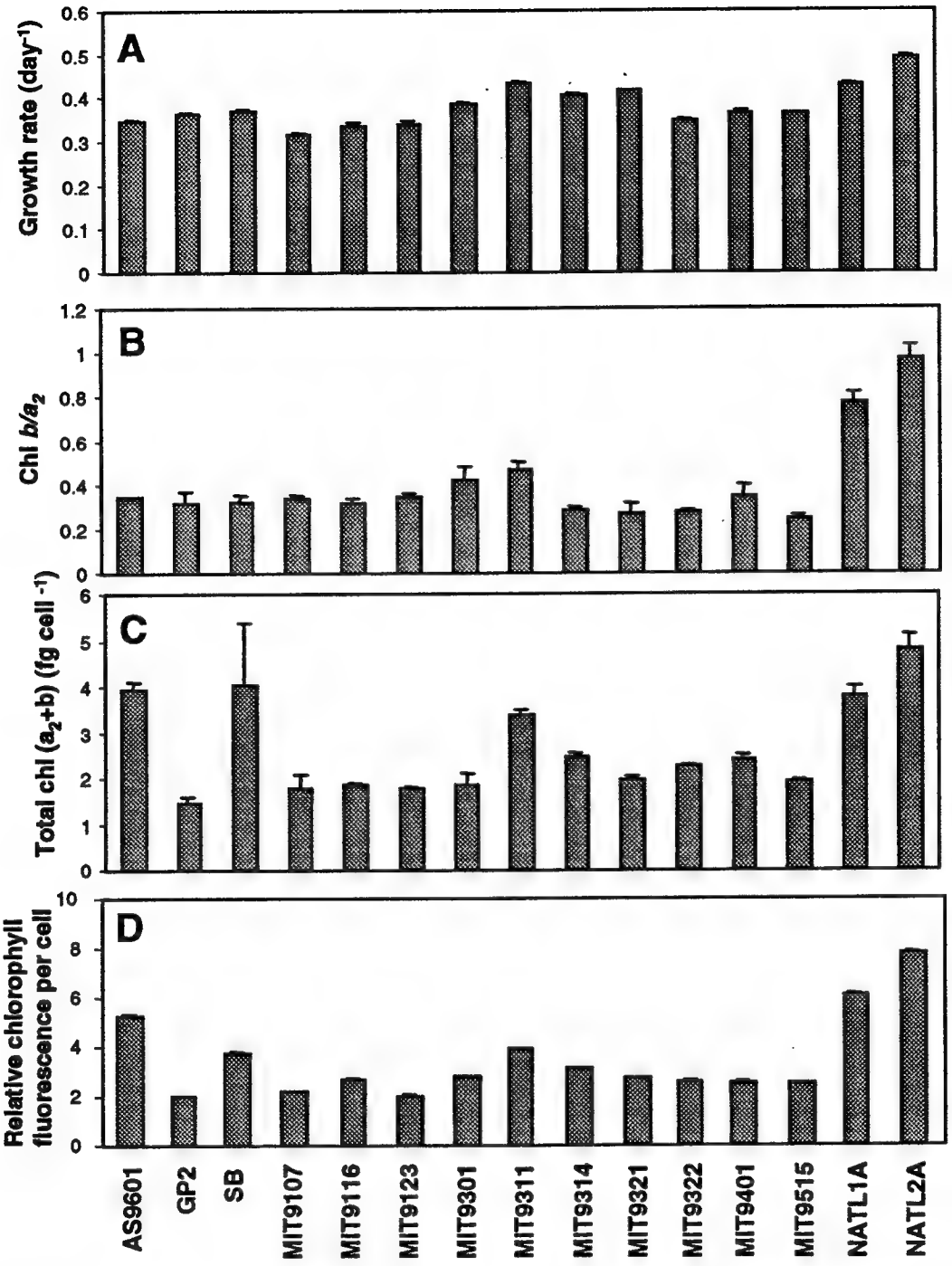


Figure 1. Growth rate, pigment and fluorescence properties of fifteen isolates of *Prochlorococcus* grown at 18 µmol Q m^{-2} s^{-1}. Values are means (± 1SE) of triplicate cultures. **(A)** Growth rate. **(B)** Chlorophyll b/a_2 ratio. **(C)** Total cellular chlorophyll ($a_2 + b$). **(D)** Mean chlorophyll (red) fluorescence per cell, relative to standard beads, derived from flow cytometry.

and Vaulot 1993, Veldhuis and Kraay 1993) and can reflect the coexistence of multiple ecotypes (Moore, et al. 1998). Here, the flow cytometrically derived chlorophyll fluorescence was strongly correlated with the total chlorophyll per cell (r^2=0.83) and the chl b/a_2 ratio (r^2=0.76) and weakly correlated with growth rates of the isolates (Fig. 2).

Sequences of ITS. Sequences of the 16S 23S rDNA internal transcribed spacer region were determined for 30 strains of *Prochlorococcus* and 16 clonal strains of *Synechococcus*. Although the majority of the *Prochlorococcus* isolates have not been rendered clonal, PCR amplifications yielded single band products and few sequence ambiguities were observed. All of the strains examined contained genes encoding two tRNAs, for Isoleucine and Alanine, as has been observed in freshwater *Synechococcus* PCC6803 (Tomioka and Sugiura 1984).

Sequencing of the 46 strains resulted in 34 unique ITS sequences. *Prochlorococcus* strains SS120, SS2, SS35, SS51 are clonal derivatives of the primary culture SARG, and all 5 strains had identical sequences. *Prochlorococcus* strain MED4Ax also had an identical sequence to its parent strain MED4. *Prochlorococcus* strains MIT9107, MIT9116, MIT9123, co-isolates from the same water sample in the South Pacific, were identical to each other, as were co-isolates MIT9312 and MIT9311 from the Gulf Stream. The *Prochlorococcus* strains MIT9321, MIT9322 and MIT9401 also possessed identical ITS sequences, although MIT9401was isolated from the Sargasso Sea while the other two are from the Equatorial Pacific. *Synechococcus* WH7805, WH8008, and WH8018,

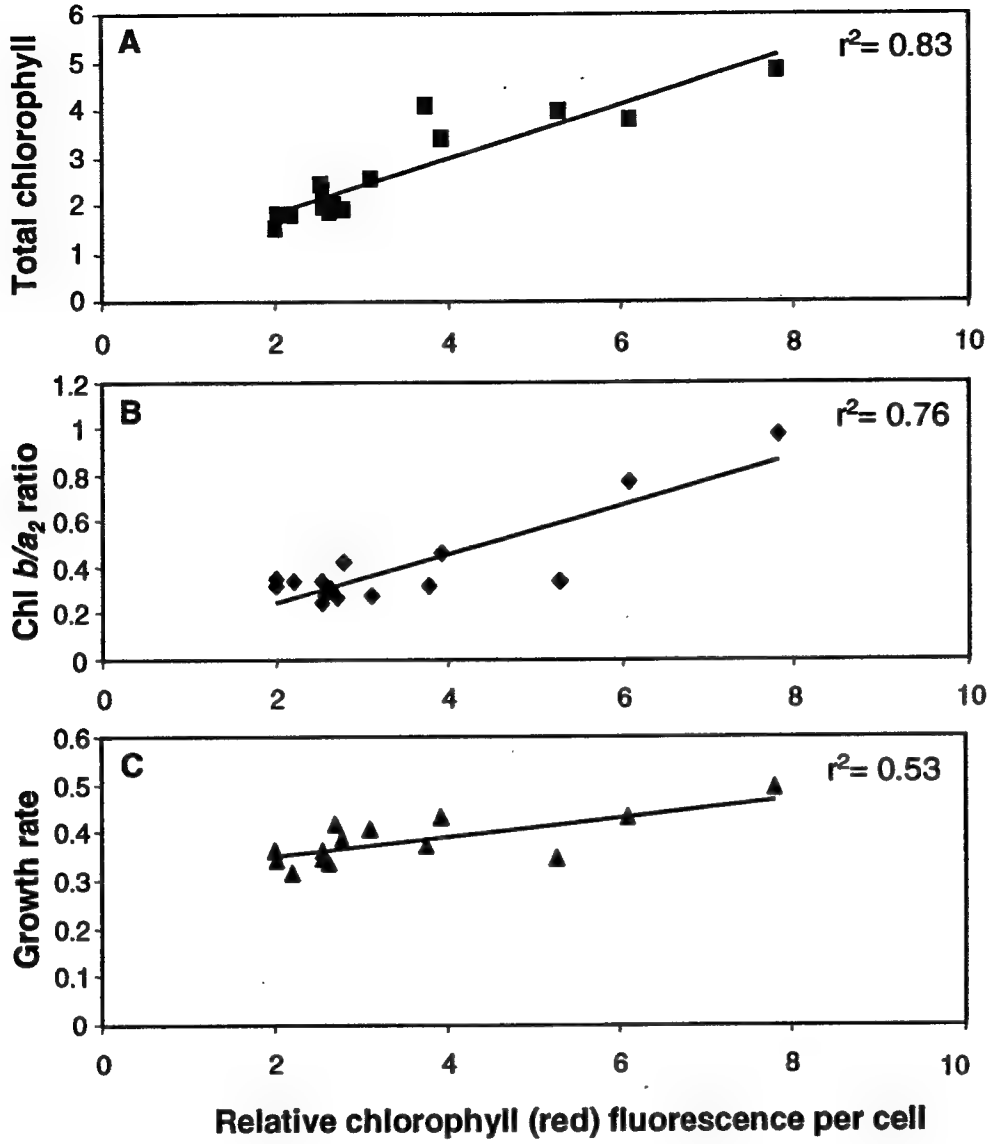


Figure 2. Correlation of mean chlorophyll (red) fluorescence per cell with other physiological properties for 15 isolates of *Prochlorococcus* (**A**) Total chlorophyll (**B**) chl b/a_2 ratio (**C**) growth rate

isolated from the Sargasso Sea, the Gulf of Mexico and Woods Hole respectively, also possess identical ITS sequences. For the remainder of this paper results will be presented only for the 34 unique sequences, with SS120, MED4, MIT9107, MIT9312, MIT9401, and WH7805 used to represent the six sets of identical strains.

Marked differences in the length of the ITS were observed among the 34 unique sequences (Fig. 3). Lengths of the ITS ranged from 539 bp in *Prochlorococcus* MIT9314 to 1017 bp in *Synechococcus* PCC6307. The length differences were strongly correlated with ecotype. All of the low B/A *Prochlorococcus* have ITS regions ranging in length from 541-550 bp. The high B/A *Prochlorococcus* have much longer ITS sequences, and there is a larger range of lengths among them. NATL1A, NATL2A, SS120 and MIT9211 ITS sequences range in length from 634 to 695 bp, while MIT9303 and MIT9313 have spacers of 833 and 831 bp respectively. Marine *Synechococcus* strains range from 749 bp (WH8017) to 812 bp (WH8103). The majority of the length difference is in the 3' end of the spacer ($tRNA^{Ala}$ - 23S spacer) which ranges from 254-528 bp across the strains examined (Fig. 3).

Substantial differences also exist in the %G+C content of the ITS (Fig. 4). Low B/A *Prochlorococcus* ITS sequences had the lowest %G+C content (37-39%). Again, the high B/A *Prochlorococcus* spanned a range of values. The %G+C content of sequences from NATL1A, NATL2A, SS120 and MIT9211 was quite similar to that of the low B/A isolates. However MIT9303 and MIT9313 had higher %G+C contents of 44 and 45%.

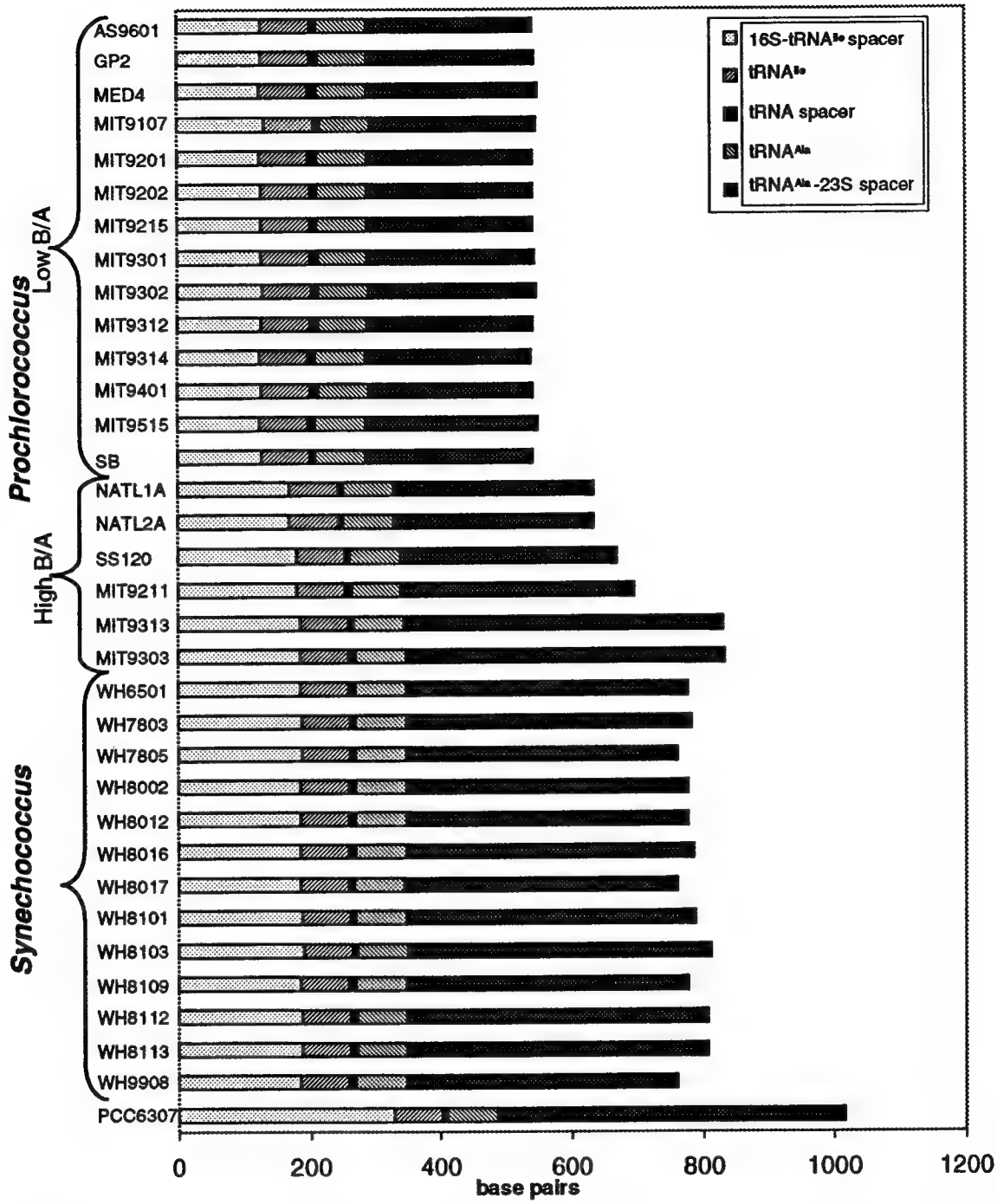


Figure 3. Length of the internal transcribed spacer region (ITS) between the 16S and 23S rDNAs in *Prochlorococcus* and *Synechococcus* isolates. Low B/A *Prochlorococcus* isolates all have short spacers, while high B/A isolates have longer spacer lengths and are not as similar to one another. The freshwater strain PCC6307 has a much longer spacer than the marine *Synechococcus* and *Prochlorococcus* isolates.

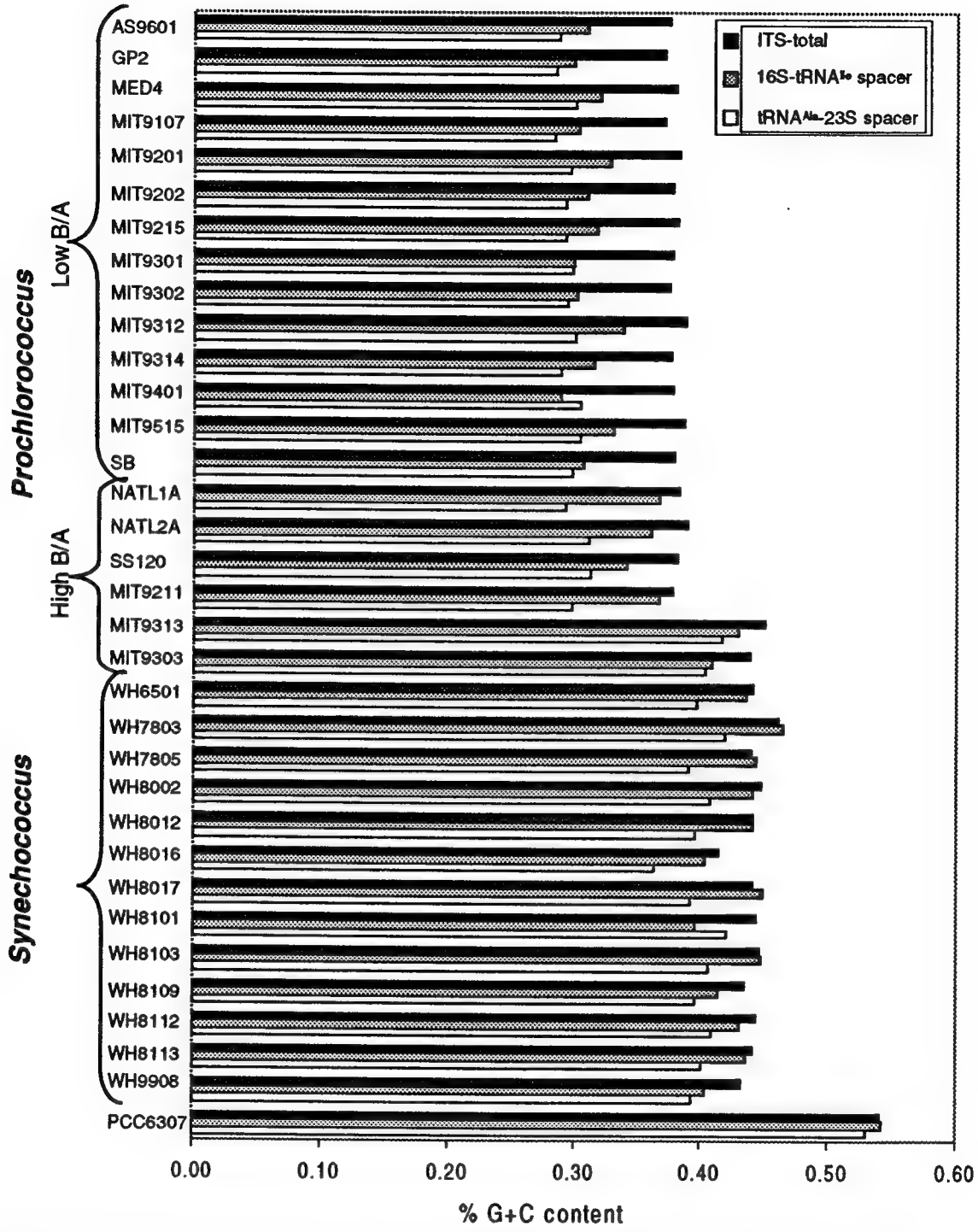


Figure 4. % G+C content of the 16S 23S rDNA internal transcribed spacer (ITS) region and its two most variable regions, the 16S- $tRNA^{Ile}$ spacer and the $tRNA^{Ala}$-23S spacer. Low B/A *Prochlorococcus* and 4 high B/A *Prochlorococcus* isolates have a lower %GC content than 2 high B/A *Prochlorococcus* and *Synechococcus* isolates.

Values for MIT9303 and MIT9313 are in the range of the majority of the marine *Synechococcus* strains (44-46 %G+C). Two *Synechococcus* strains, WH8016 and WH9908, did have lower %G+C contents (41 and 43%). As with its much longer length, the %G+C content of the ITS sequence from Cyanobium PCC6307 was markedly different (54%G+C) from all of the *Prochlorococcus* and marine *Synechococcus.* The majority of the strains had lower %G+C contents in the two spacer regions than in the ITS as a whole, presumably because of the constraints necessary to maintain secondary structure of the tRNAs. The $tRNA^{Ala}$ - 23S spacer region sequences were the lowest in %G+C content and also exhibited the widest range among the isolates (Fig. 4).

Phylogenetic relationships. Phylogenetic analyses used the 16S-$tRNA^{Ile}$ spacer because the more extreme differences in length and %G+C content (Fig. 3 and 4) made alignments difficult for the $tRNA^{Ala}$ - 23S spacer. Because the shorter length of the 16S-$tRNA^{Ile}$ spacer in the low B/A *Prochlorococcus* limited the positions available in the unambiguously aligned dataset, parallel analyses were conducted with and without the low B/A *Prochlorococcus*. Phylogenetic analyses used 103 nucleotide positions when all strains were considered and an expanded dataset of 146 nucleotide positions when only the high B/A *Prochlorococcus* and *Synechococcus* were considered (Fig. 5). Cyanobium cluster *Synechococcus* branches outside of the marine picoplankton clade in 16S rDNA analyses (Urbach, et al. 1998) and was included here as an outgroup. However, the large degree of sequence dissimilarity between PCC6307 and the marine sequences precluded confident alignments, so it was excluded from phylogenetic analyses.

Figure 3. Alignments of 16S-tRNA^{Ile} spacer region sequences in Prochlorococcus and Synechococcus isolates. Positions used in phylogenetic analyses indicated by an asterisk.

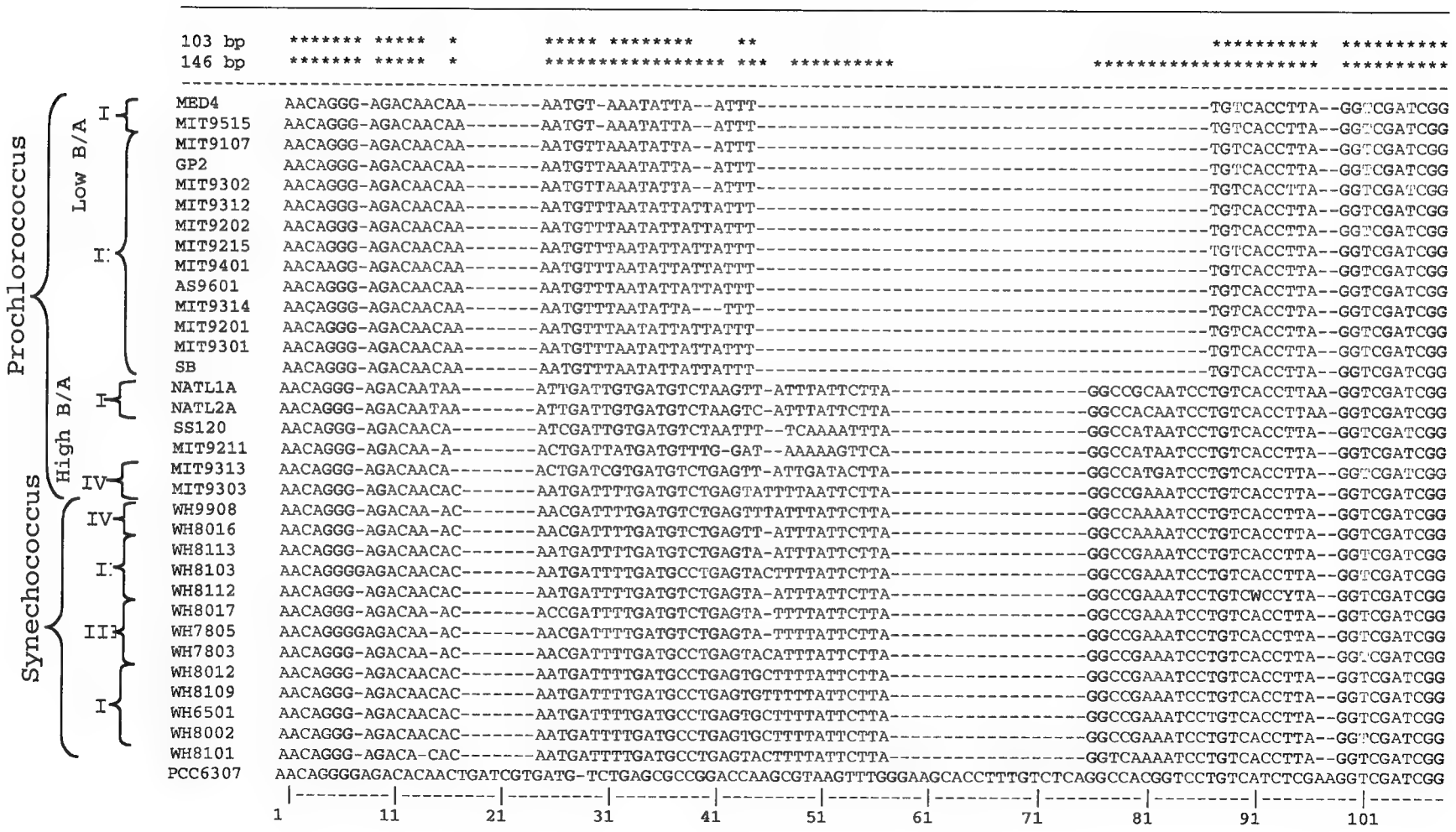


Figure 5(Continued):

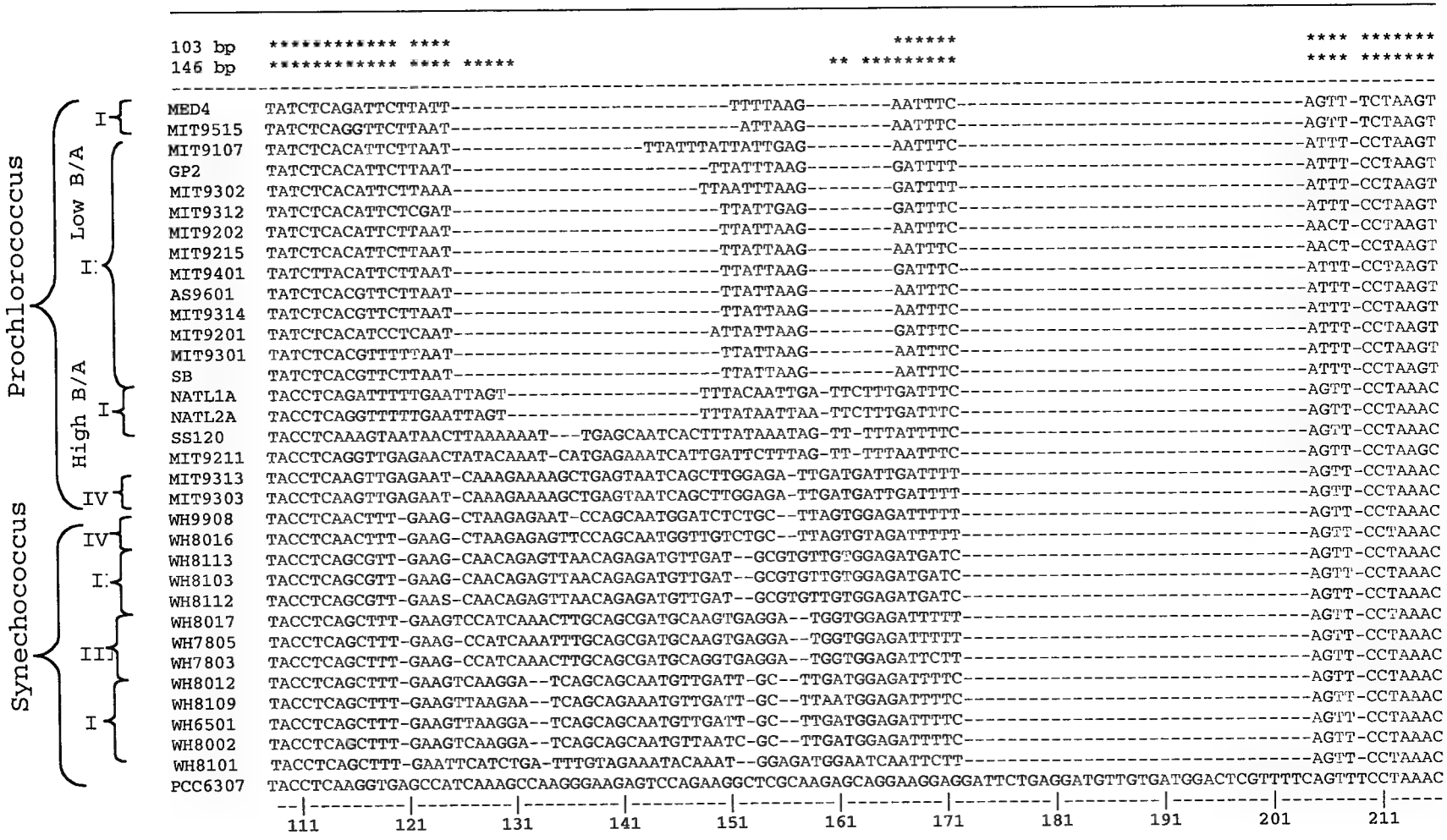
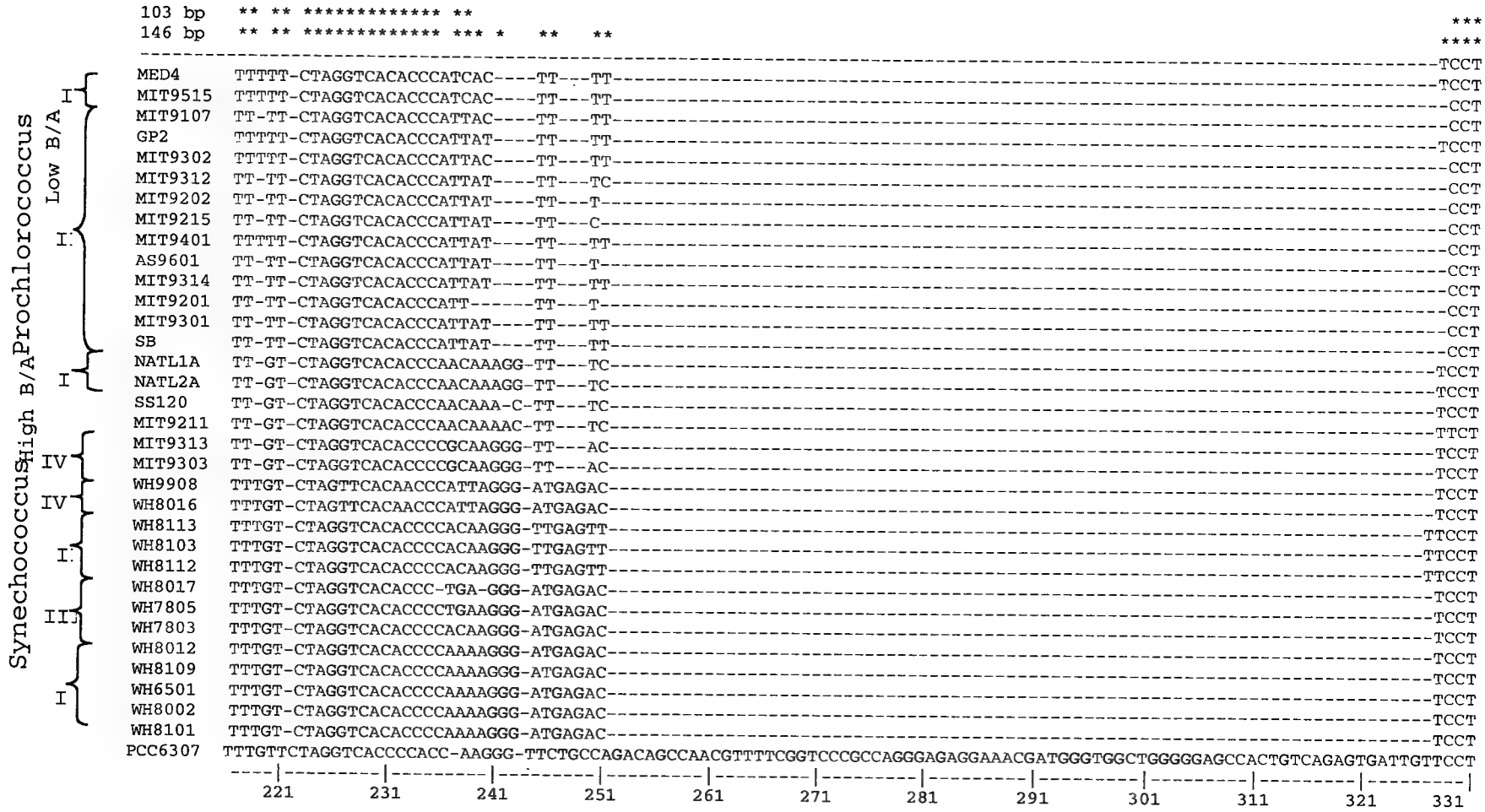


Figure 5. (Continued):



In phylogenetic analyses confined to 103 base pairs of the 16S-tRNA^{Ile} spacer, the low B/A *Prochlorococcus* form a well supported clade, which can be further subdivided into two clusters, one containing MED4 and MIT9515, and the other containing the remaining low B/A isolates (Fig. 6). This division is reflected in the amount of sequence difference between the low B/A isolates. The low B/A cluster II isolates are 93-99% similar in their 16S-tRNA^{Ile} spacer sequences, while MED4 and MIT9515 are only 89-93% similar to the low B/A cluster II isolates (Table 3).

The high B/A *Prochlorococcus* do not form a monophyletic clade, but are dispersed in four distinct lineages (Fig. 6). One cluster, designated high B/A I, contains isolates NATL1A and NATL2A which consistently branch together and share a high spacer sequence similarity (97%). Isolates SS120 and MIT9211 are much less similar to each other (82%). Although distance methods provide some support for SS120 and MIT9211 branching together, parsimony and likelihood methods do not (Fig. 6). Since neither isolate has a high sequence similarity to each other or to the other four high B/A isolates (Table 3), they have each been assigned to a different cluster (Fig. 6). Finally, isolates MIT9303 and MIT9313 make up *Prochlorococcus* high B/A cluster IV. Although support for this grouping varies somewhat depending on the method of tree inference (Fig. 6), the two strains are 95% similar in their 16S-tRNA^{Ile} spacer sequences (Table 3) and have been shown to be quite similar physiologically (Moore and Chisholm 1999). These same clusters are apparent in analyses using a larger subset (146 bp) of the 16S-tRNA^{Ile} spacer (Fig. 7).

Table 3. Fractional similarities for the 16S-tRNA^{Ile} spacer

			MED4	MIT9515	AS9601	GP2	MIT9107	MIT9201	MIT9202	MIT9215	MIT9301	MIT9302	MIT9312	MIT9401	NATL1A	NATL2A
Prochlorococcus	Low B/A I	MED4	–													
		MIT9515	0.98	–												
	Low B/A II	AS9601	0.93	0.94	–											
		GP2	0.93	0.93	0.97	–										
		MIT9107	0.94	0.93	0.97	0.97	–									
		MIT9201	0.91	0.92	0.97	0.97	0.95	–								
		MIT9202	0.93	0.93	0.98	0.96	0.96	0.96	–							
		MIT9215	0.92	0.92	0.97	0.95	0.95	0.95	0.99	–						
		MIT9301	0.92	0.93	0.99	0.96	0.96	0.96	0.97	0.96	–					
		MIT9302	0.92	0.91	0.94	0.98	0.96	0.94	0.94	0.93	0.94	–				
		MIT9312	0.89	0.89	0.96	0.95	0.95	0.98	0.95	0.94	0.94	0.93	–			
		MIT9401	0.91	0.91	0.97	0.97	0.95	0.97	0.96	0.95	0.96	0.94	0.95	–		
	High B/A	NATL1A	0.78	0.77	0.75	0.75	0.75	0.74	0.75	0.74	0.76	0.77	0.76	0.74	–	
		NATL2A	0.78	0.79	0.76	0.75	0.75	0.73	0.74	0.74	0.76	0.77	0.75	0.73	0.97	–
		SS120	0.76	0.76	0.74	0.72	0.71	0.75	0.74	0.73	0.73	0.74	0.75	0.72	0.84	0.85
		MIT9211	0.73	0.76	0.73	0.71	0.70	0.73	0.72	0.71	0.73	0.71	0.73	0.70	0.78	0.79
		MIT9303	0.68	0.69	0.67	0.67	0.63	0.68	0.65	0.65	0.67	0.67	0.68	0.65	0.78	0.78
		MIT9313	0.70	0.70	0.69	0.68	0.65	0.70	0.67	0.67	0.69	0.69	0.70	0.67	0.78	0.78
Synechococcus	I	WH6501	0.69	0.70	0.67	0.68	0.67	0.69	0.66	0.66	0.67	0.68	0.69	0.66	0.77	0.76
		WH8002	0.68	0.69	0.67	0.67	0.66	0.68	0.66	0.66	0.66	0.68	0.69	0.66	0.77	0.76
		WH8012	0.69	0.70	0.67	0.68	0.67	0.69	0.66	0.66	0.66	0.68	0.69	0.66	0.77	0.76
		WH8109	0.70	0.72	0.70	0.68	0.69	0.69	0.69	0.69	0.69	0.68	0.70	0.67	0.77	0.77
	II	WH8103	0.70	0.72	0.67	0.67	0.67	0.67	0.66	0.65	0.67	0.66	0.67	0.66	0.77	0.76
		WH8113	0.71	0.73	0.68	0.68	0.68	0.69	0.67	0.66	0.68	0.67	0.68	0.67	0.78	0.77
	III	WH7803	0.69	0.70	0.67	0.69	0.65	0.68	0.66	0.66	0.67	0.68	0.69	0.66	0.76	0.75
		WH7805	0.70	0.71	0.68	0.69	0.64	0.69	0.67	0.67	0.68	0.68	0.69	0.67	0.75	0.74
		WH8017	0.71	0.72	0.69	0.71	0.66	0.69	0.68	0.68	0.69	0.70	0.70	0.68	0.75	0.75
	IV	WH8016	0.68	0.68	0.67	0.67	0.65	0.68	0.66	0.66	0.67	0.67	0.68	0.64	0.77	0.77
		WH9908	0.67	0.68	0.67	0.66	0.64	0.68	0.66	0.66	0.67	0.67	0.68	0.64	0.75	0.76
		WH8101	0.69	0.71	0.68	0.70	0.66	0.68	0.67	0.67	0.67	0.70	0.69	0.67	0.77	0.78

Prochlorococcus strains AS9601, MIT9314 and SB have identical 16S-tRNA^{Ile} spacers, as do *Synechococcus* strains WH8112 and WH8113

Table 3 continued

	SS120	MIT9211	MIT9303	MIT9313	WH6501	WH8002	WH8012	WH8109	WH8103	WH8113	WH7803	WH7805	WH8017	WH8016	WH9908
MED4															
MIT9515															
AS9601															
GP2															
MIT9107															
MIT9201															
MIT9202															
MIT9215															
MIT9301															
MIT9302															
MIT9312															
MIT9401															
NATL1A															
NATL2A															
SS120	–														
MIT9211	0.82	–													
MIT9303	0.76	0.74	–												
MIT9313	0.78	0.74	0.95	–											
WH6501	0.71	0.71	0.81	0.77	–										
WH8002	0.70	0.70	0.82	0.78	0.98	–									
WH8012	0.70	0.70	0.82	0.78	0.99	0.99	–								
WH8109	0.71	0.71	0.81	0.76	0.98	0.96	0.97	–							
WH8103	0.71	0.71	0.76	0.71	0.89	0.88	0.89	0.89	–						
WH8113	0.72	0.72	0.76	0.72	0.88	0.87	0.88	0.88	0.99	–					
WH7803	0.72	0.68	0.80	0.76	0.89	0.89	0.89	0.87	0.85	0.85	–				
WH7805	0.72	0.69	0.81	0.76	0.88	0.89	0.89	0.87	0.84	0.84	0.94	–			
WH8017	0.71	0.68	0.81	0.77	0.88	0.89	0.89	0.87	0.83	0.83	0.94	0.99	–		
WH8016	0.75	0.71	0.77	0.75	0.85	0.85	0.85	0.86	0.81	0.82	0.86	0.86	0.84	–	
WH9908	0.75	0.70	0.78	0.76	0.85	0.85	0.85	0.86	0.81	0.82	0.86	0.87	0.85	0.98	–
WH8101	0.70	0.67	0.79	0.76	0.85	0.86	0.85	0.84	0.80	0.79	0.88	0.88	0.87	0.80	0.80

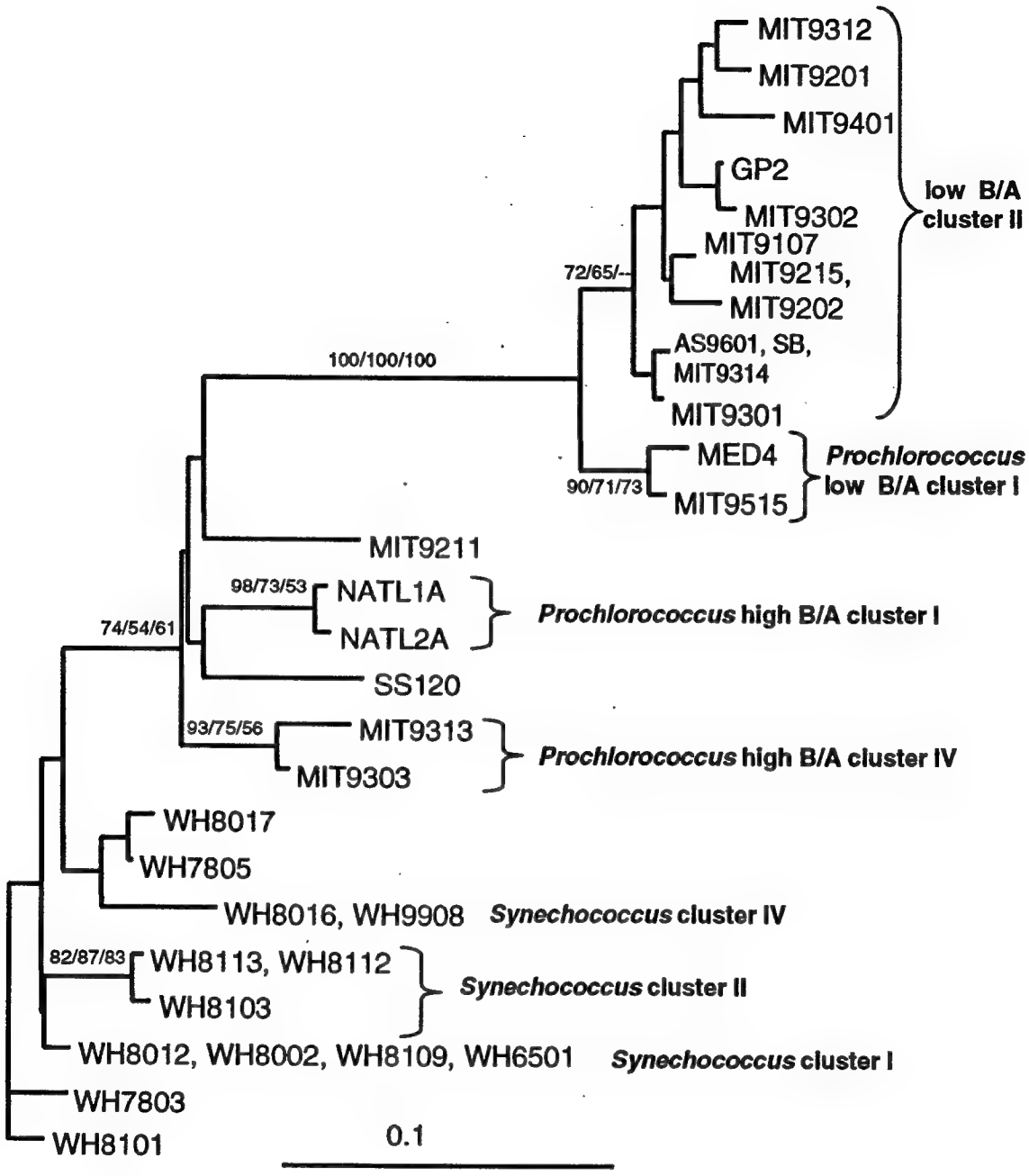


Figure 6. Evolutionary relationships of *Prochlorococcus* and *Synechococcus* isolates inferred using 103 positions of the 16S- $tRNA^{Ile}$ spacer. The phylogenetic framework was determined using paralinear distances (logdet) and minimum evolution as the objective function. Trees inferred using HKY85 distances resulted in identical clusters with essentially similar branching orders. Bootstrap values from [distance/maximum parsimony/maximum likelihood] analyses are listed to the left of each node, with values less than 50 not shown. The tree is arbitrarily rooted with WH8101.

The expanded subset of the 16S-tRNA^{Ile} spacer allowed the designation of strains of marine A *Synechococcus* are in four different clusters (Fig. 7). One, designated cluster I, consists of the low PUB strains WH6501, WH8002, WH8012 and WH8109 (Fig. 7). These strains are very similar in their spacer sequence (96-99%) (Table 3). A second well supported cluster (II) consists of the motile high PUB strains WH8103, WH8112, and WH8113. These strains also show very high sequence similarity (99%). A third cluster, more physiologically and genetically divergent than the first two, consists of the strain WH7805, which lacks PUB (and has an identical sequence to strains WH8008 and WH8018, which also lack PUB) and the low PUB strains WH7803 and WH8017. Strains WH7805 and WH8017 differ by only 1% over the 16S-tRNA^{Ile} spacer, while WH7803 is 94 % similar to these two strains. Finally, *Synechococcus* cluster IV contains strains WH9908 and WH8016 (only 2% different), which form a well supported cluster, although little can be said about the physiological characteristics which may define this group. WH8016 is a low PUB strain, but pigment data is not yet available for the psychrophilic strain WH9908 which was isolated in Woods Hole when water temperatures were below 10° and has been maintained at 15°C and shows no growth at 23°C (M. Sullivan, personal communication).

Although the expanded dataset of 146 positions was used to designate strain clusters of high B/A *Prochlorococcus* and *Synechococcus* because it offers greater resolution, some of these clusters are also apparent on the tree constructed using only 103 positions (Fig. 6 and 7). For example high B/A clusters I and IV are well supported, as is the monophyly

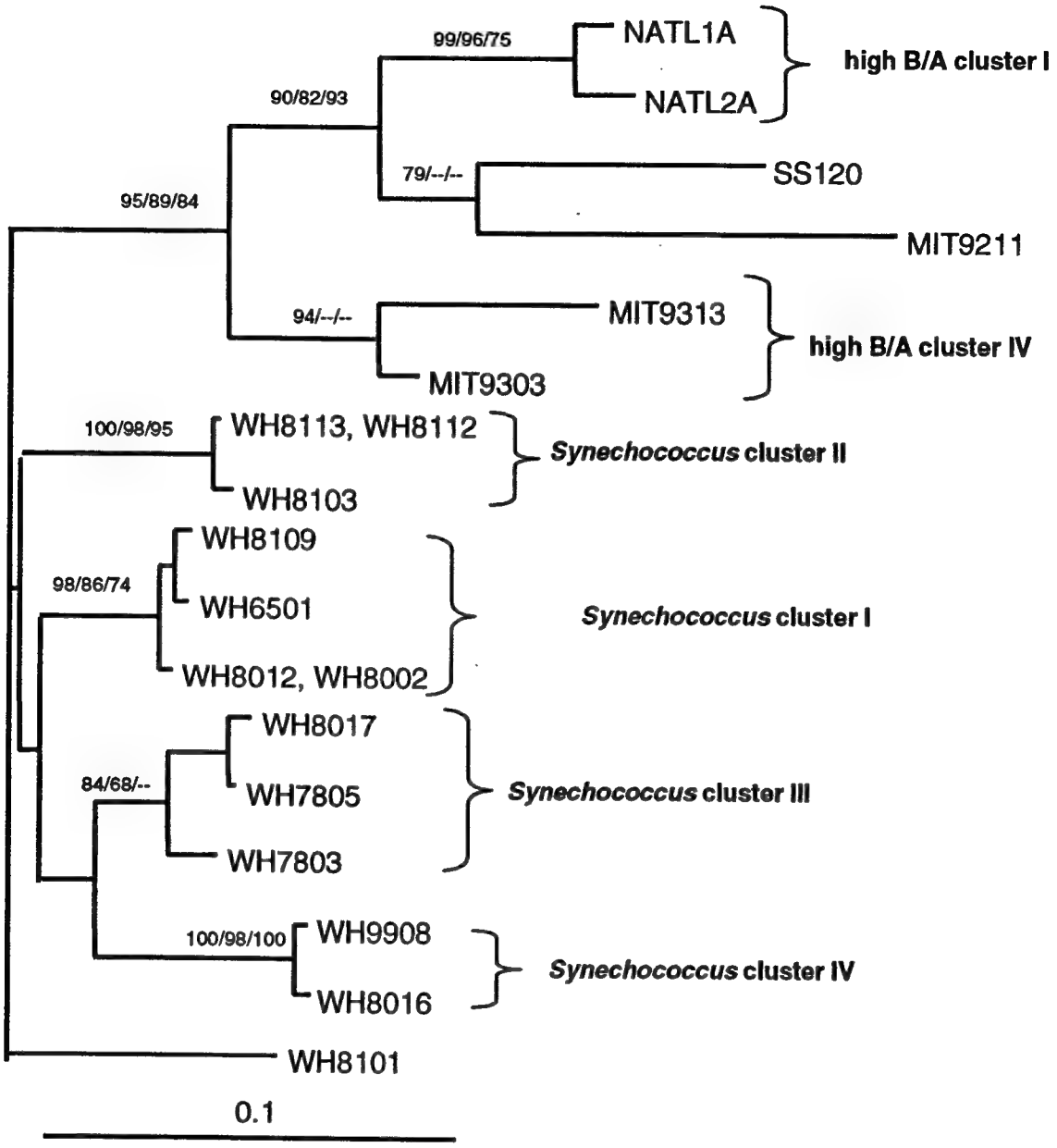


Figure 7. Evolutionary relationships of high B/A *Prochlorococcus* and *Synechococcus* isolates inferred using 146 positions of the 16S- $tRNA^{Ile}$ spacer. The phylogenetic framework was determined using paralinear distances (logdet) and minimum evolution as the objective function. Trees inferred using HKY85 distances resulted in identical clusters with essentially similar branching orders. Bootstrap values from [distance/maximum parsimony/maximum likelihood] analyses are listed to the left of each node, with values less than 50 not shown. The tree is rooted with WH8101.

of the *Prochlorococcus* clade as a whole, although it is less well supported by bootstrap values with the smaller dataset. *Synechococcus* cluster II is also supported with the smaller dataset, while the two strains in *Synechococcus* cluster IV are identical to one another over these 103bp, as are the four strains in *Synechococcus* cluster I.

The marine B *Synechococcus* isolate WH8101 is 80-88% similar in the sequence of the 16S-tRNA^{Ile} spacer to the marine A isolates, approximately as distinct as each of the marine A *Synechococcus* clusters are to one another (Table 3). This supports the suggestion that the lineages of marine A *Synechococcus* are of equivalent taxonomic rank to the division between marine A and B clusters (Urbach, et al. 1998).

DISCUSSION

The higher variability of the ITS compared to the 16S rDNA has enabled the identification of sequence similarity clusters among the closely related isolates of *Prochlorococcus* and *Synechococcus*. Given the phylogenetic divergence of the high B/A isolates into four clusters, the question arises whether the designation high and low B/A ecotypes is still valid for *Prochlorococcus*. Recently, Rippka et al. (2000) have proposed a subspecies *pastoris* to describe low B/A *Prochlorococcus*. The proposed type strain PCC9511 has a 16S rDNA sequence identical to the well characterized low B/A isolate MED4 (Rippka, et al. 2000). There are undoubtedly more taxonomic distinctions to be made, as each of the high B/A clusters is a lineage of equal rank to the low B/A ecotype. However, in the absence of more striking ecologically relevant physiological

differences among the high B/A isolates, the designations high and low B/A ecotype remain useful. They accurately describe culture physiologies which are correlated with phylogenetic relationships and are convenient terms, to be used with the caveat that the high B/A ecotype consists of a wider range of physiologies and genetic diversity than the low B/A ecotype.

Physiological characterization of 15 *Prochlorococcus* isolates by measuring their chl b/a_2 ratio at one light level enabled the assignment of strains NATL1A and NATL2A to the high B/A ecotype and the remaining thirteen isolates to the low B/A ecotype. These designations are supported by the phylogenetic relationships of the isolates using the ITS sequences. The chl b/a_2 ratios of NATL1A and NATL2A are on the low end of values observed for high B/A isolates of *Prochlorococcus*. However, values determined here using a spectrophotometer may be underestimates of chl b/a_2 compared to values derived from HPLC methods as employed by Moore and Chisholm (1999). Others have found that estimates of chl b/a_2 for the same culture measured by spectrofluorometry were 70% of those determined using HPLC (Morel, et al. 1993). The chl b/a_2 ratio of NATL1A is similar to the HPLC-determined value of 0.95 (when grown under 7.5 μmol Q m^{-2} s^{-1} white light) first reported for this isolate (Partensky, et al. 1993). Additionally, the range of chl b/a_2 ratios for the four high B/A isolates is broader than for the low B/A isolates. SS120 and MIT9211 have much higher ratios (0.7 - 2.5) over all irradiances than MIT9303 and MIT9313 (0.5 - 1.3) (Moore and Chisholm 1999). This physiological

diversity is consistent with greater genetic diversity among high B/A isolates seen previously using 16S rDNA sequences (Rocap, et al. 1999) and here with ITS sequences.

With the addition of physiological data for these 15 isolates, it is clear that the majority of the *Prochlorococcus* isolates in culture collections are of the low B/A ecotype (19 as compared to only 6 high B/A isolates, of the 25 physiologically characterized strains in this study). This is in spite of concentrated efforts to culture additional high B/A isolates by sampling at depth, where they are presumed to predominate, and then mimicking the *in situ* light regime. Although it is possible that low B/A *Prochlorococcus* are actually numerically dominant in many parts of the ocean and our collection of isolates represents their natural frequencies, this is seems unlikely given the potential sources of isolation bias, which could select for low B/A strains. For example, filtration steps necessary to eliminate the larger *Synechococcus* could also preferentially remove larger *Prochlorococcus* similar to the high B/A cluster IV isolates MIT9303 and MIT9313, which exhibit much higher forward angle light scatter (a parameter related to cell size) than other isolates of *Prochlorococcus* (Moore 1997). It is noteworthy that both of these isolates were obtained by flow cytometric sorting, not filtration. Additionally, culture media may favor low B/A *Prochlorococcus* (possibly because of differing trace metal requirements or tolerances) and they may overtake mixed cultures.

The observation of multiple isolates with identical ITS sequences underscores the cosmopolitan nature of these genera and is consistent with previous studies which

identified *Prochlorococcus* isolates with identical 16S rDNA sequences (MIT9202, SB and TATL2) from disparate regions of the worlds oceans (Urbach, et al. 1998, Rocap, et al. 1999). At the more variable ITS, *Prochlorococcus* MIT9321 and MIT9322, isolated within days of each other in the Equatorial Pacific, are also identical to MIT9401 from the Sargasso Sea. Similarly *Synechococcus* strains WH8008, WH8018 and WH7805 isolated from the Gulf of Mexico, Woods Hole, and the Sargasso Sea respectively, have identical ITS sequences.

Of course, the sensitivity of the PCR to even minute amounts of DNA means that caution is warranted in deeming sequences identical, especially in a situation where the same set of primers is used repeatedly on different templates. However, several lines of evidence suggest that the identical sequences are features of the isolates and not due to contamination. First, because the genomic DNA from cultures being investigated was in plentiful supply, PCR reactions employed generous amounts of template DNA. Second, reactions lacking template were always run as a control and did not result in amplification products as judged by agarose gels. Third, many of the identical sequences were prepared several months apart from one another and other cultures were amplified in the interim which yielded different sequences, suggesting there was not a rampant contamination problem. Finally, it is noteworthy that in no case did a culture yield a sequence which was incongruent with its physiology, as might be expected to occur if the identical sequences were the result of chance contamination.

The monophyly of low B/A *Prochlorococcus* demonstrated here with the ITS (Fig. 6) is consistent with previous results from 16S rDNA and other protein coding genes (Urbach, et al. 1998). The subdivision of the low B/A ecotype is also consistent with 16S rDNA analyses in which a well supported branching pattern separates MED4 from other low B/A isolates (Rocap, et al. 1999). Although no significant physiological differences have been observed between isolates of these two clusters, they may have differential distributions in the ocean, supporting the contention that sequence similarity clusters as defined by the ITS correspond to ecologically distinct populations. In surface waters in the North Atlantic MED4-like *Prochlorococcus* sequences were detected by using probes to amplified 16S rDNA while the low B/A cluster II could not be detected (West and Scanlan 1999).

High B/A *Prochlorococcus* are distributed in four clusters, each only about 80% similar to one another. This greater genetic divergence among the high B/A *Prochlorococcus* is consistent with the larger differences observed in their physiologies. For example, in addition to the higher chl b/a_2 ratio of SS120 and MIT9211 and the larger cell size of MIT9303 and MIT9313 already described (Moore 1997, Moore and Chisholm 1999), MIT9303 and MIT9313 also possess an unidentified pigment, possibly an epoxide of isocryptoxanthin (Moore 1997). Further, when grown under high light MIT9303 and MIT9313 have higher cell-specific light harvesting efficiencies and maximum photosynthetic rates than SS120 and MIT9211 (Moore and Chisholm 1999). As with the two low B/A clusters, there is some evidence that the high B/A clusters may have

different distributions in the oceans. In a North Atlantic water column sequences corresponding to the NATL1A/NATL2A type (high B/A cluster I) were detected by probing PCR amplified 16S rDNA, while binding to probes specific for SS120 and MIT9303 was not detected (West and Scanlan 1999).

Marine A *Synechococcus* encompass at least four sequence similarity clusters. One of these, cluster II can be associated with a physiological trait as all of its member strains are capable of swimming. Motile *Synechococcus* very likely occupy a unique niche because the ability to migrate towards patches of nitrogenous compounds (e.g., on marine snow) allows them to access nutrient sources other *Synechococcus* and *Prochlorococcus* cannot (Willey and Waterbury 1989). This cluster is also consistent with analyses using *rpoC1* in which motile isolates with a range of PUB types form a single clade (Toledo, et al. 1999). Analyses using *rpoC1* also affiliate cluster III isolates WH7803 and WH7805 (Toledo and Palenik 1997). Unfortunately it is not yet possible to completely compare the clusters identified using the ITS with those delineated by *rpoC1* because the majority of the strains used in the two studies are different. More strains need to be examined using both *rpoC1* and ITS sequences to make definitive conclusions about correlations of sequence clusters and physiological (particularly pigment) type.

The branching order of the low B/A *Prochlorococcus* clade with respect to the high B/A clusters and *Synechococcus* suggests that the low B/A clade is more recently arisen. If the ability to synthesize divinyl chlorophylls *a* and *b* has been acquired only once (i.e. if

the root of the tree is outside the *Prochlorococcus* clade, which is reasonable to assume) then the general possession of divinyl chlorophyll *a* and *b* is a primitive state shared by all *Prochlorococcus,* and the low B/A clade is a derived state. This is consistent with a possible evolutionary scenario in which a phycobilisome containing ancestor acquired the ability to synthesize divinyl chlorophylls *a* and *b*, allowing it to colonize the deep euphotic zone, where competition for nutrients is reduced, as no other photosynthetic organisms can thrive there. Then later modifications, including the ability to survive in higher light and higher copper concentrations characteristic of surface waters, enabled the low B/A clade to be more competitive throughout the water column.

CONCLUSIONS

Using the sequences of the 16S 23S rDNA internal transcribed spacer region we have successfully delineated these two genera of marine cyanobacteria into strain clusters which likely represent ecological units. For some of the clusters, for example the low B/A *Prochlorococcus,* some inferences can be made about the nature of the niche these organisms are occupying (here, high light surface waters). In other cases, as with some of the *Synechococcus* clusters, there is not an obvious hypothesis to explain the different niches these lineages may be occupying. The contribution of this work is in identifying the sequence similarity clusters which potentially correspond to ecologically distinct units. This allows for more informed selection of strains in laboratory experiments, as representative strains from each lineage can now be employed in further physiological studies to discover the features of each cluster which may lead to niche differentiation.

The examination of sequences directly from the environment, thus avoiding culturing biases, has been widespread in recent years and has given new insights into the diversity of environmental bacteria (Giovannoni, et al. 1990, Ward, et al. 1990) including marine cyanobacteria (Palenik 1994, Ferris and Palenik 1998, Urbach and Chisholm 1998). Sequence clusters identified by the ITS are excellent candidates for direct sequence diversity studies of field populations of marine cyanobacteria because the ITS is variable enough to differentiate the ecotypes unambiguously. Further, the sequence data presented here allows the design of oligonucleotides specific for each cluster which can be used as probes in quantitative or *in situ* hybridizations (Giovannoni, et al. 1988, DeLong, et al. 1989) or primers in quantitative PCR (Lee, et al. 1996, Tay, et al. 2000). Using these tools, the distribution and relative abundances of each cluster can be studied under different environmental conditions (Field, et al. 1997), ultimately providing a better understanding of the forces affecting the evolution and population dynamics of this globally successful clade of cyanobacteria.

ACKNOWLEDGEMENTS

This work was supported by NASA and NSF. I thank John Waterbury for cultures of *Synechococcus* and Dan Distel for ITS sequences of *Synechococcus* cultures PCC6307, WH8103, WH8112, WH8012, WH8008, WH8018, WH8017, WH7805 and WH8101. Further I thank Fred Partensky for *Prochlorococcus* cultures NATL1A and NATL2A, Rob Olson and Alexi Shalapyonok for *Prochlorococcus* culture AS9601, A. Shimada for *Prochlorococcus* strains GP2 and SB, Mak Saito for *Prochlorococcus* strain MED4Ax, and Matt Sullivan for *Synechococcus* strain WH9908. I also thank Nathan Ahlgren for sequencing assistance, and Lisa Moore for helpful discussions.

REFERENCES

Apirion, D. and A. Miczak (1993) RNA Processing in Prokaryotic Cells. *BioEssays* **15**:113-120

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl (1992) Short Protocols in Molecular Biology. John Wiley & Sons, New York

Barry, T., G. Colleran, M. Glennon, L.R. Dunican and F. Gannon (1991) The 16S/23S Ribosomal Spacer Region as a Target for DNA Probes to Identify Eubacteria. *PCR Methods and Applications* **1**:51-56

Campbell, L. and D. Vaulot (1993) Photosynthetic picoplankton community structure in the subtropical north Pacific ocean near Hawaii (station ALOHA). *Deep Sea Research* **40**:2043-2060

Chisholm, S.W., S.L. Frankel, R. Goericke, R.J. Olson, B. Palenik, J.B. Waterbury, L. West-Johnsrud and E.R. Zettler (1992) *Prochlorococcus marinus nov. gen. nov. sp.*: an oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Archives of Microbiology* **157**:297-300

Chisholm, S.W., R.J. Olson, E.R. Zettler, R. Goericke, J. Waterbury and N. Welschmeyer (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**:340-343

Christensen, H., K. Jorgensen and J.E. Olsen (1999) Differentiation of *Campylobacter coli* and *C. jejuni* by length and DNA sequence of the 16S-23S rRNA internal spacer region. *Microbiology* **145**:99-105

DeLong, E.F., G.S. Wickham and N.R. Pace (1989) Phylogenetic stains: Ribosomal RNA-based probes for the identification of single cells. *Science* **243**:1360-1363

Dusenberry, J.A. and S.L. Frankel (1994) Increasing the sensitivity of a FACScan flow cytometer to study oceanic picoplankton. *Limnology and Oceanography* **39**:206-209

Ferris, M.J. and B. Palenik (1998) Niche adaptation in ocean cyanobacteria. *Nature* **396**:226-228

Field, K.G., D. Gordon, T. Wright, M. Rappè, E. Urbach, K. Vergin and S.J. Giovannoni (1997) Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Applied and Environmental Microbiology* **63**:63-70

Giovannoni, S.J., T.B. Britschgi, C.L. Moyer and K.G. Field (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**:60-63

Giovannoni, S.J., E.F. DeLong, G.J. Olsen and N.R. Pace (1988) Phylogenetic Group-specific oligodeoxynucleotide probes for identification of single Microbial cells. *Journal of Bacteriology* **270**:720-726

Goericke, R. and N.A. Welschmeyer (1993) The Marine Prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. *Deep Sea Research* **40**:2283-2294

Gürtler, V. and V. Stanisich (1996) New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology* **142**:3-16

Kirk, J.T.O. (1994) Light and Photosynthesis in Aquatic Ecosystems. Cambridge University Press, Cambridge

Leblond-Bourget, N., H. Philippe, I. Mangin and B. Decaris (1996) 16S rRNA and 16S to 23S Internal transcribed spacer sequence analyses reveal Inter- and Intraspecific *Bifidobacterium* Phylogeny. *International Journal of Systematic Bacteriology* **46**:102-111

Lee, S.-Y., J. Bollinger, D. Bezdicek and A. Ogram (1996) Estimation of the abundance of an uncultured soil bacterial strain by a competitive quantitative PCR method. *Applied and Environmental Microbiology* **62**:3787-3793

Li, W.K.W. (1995) Composition of ultraphytoplankton in the central North Atlantic. *Marine Ecology Progress Series* **122**:1-8

Liu, H., H.A. Nolla and L. Campbell (1997) *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. *Aquatic Microbial Ecology* **12**:39-47

Maidak, B.L., et al. (1999) A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Research* **27**:171-173

Mann, E.L. (2000) Trace Metals and the Ecology of Marine Cyanobacteria. Ph.D. Thesis. Biology, MIT and WHOI, Cambridge, MA

Moore, L.R. (1997) Physiological ecology of *Prochlorococcus*: A comparison of isolates from diverse oceanic regimes. Ph.D. Thesis. Civil and Environmental Engineering, MIT, Cambridge

Moore, L.R. and S.W. Chisholm (1999) Photophysiology of the Marine Cyanobacterium *Prochlorococcus:* Ecotypic differences among cultured isolates. *Limnology and Oceanography* **44**:628-638

Moore, L.R., R.E. Goericke and S.W. Chisholm (1995) Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Marine Ecology Progress Series* **116**:259-275

Moore, L.R., G. Rocap and S.W. Chisholm (1998) Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* Ecotypes. *Nature* **393**:464-467

Morel, A., Y.-H. Ahn, F. Partensky, D. Vaulot and C. Hervè (1993) *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. *Journal of Marine Research* **51**:617-649

Ong, L.J. and A.N. Glazer (1991) Phycoerythrins of Marine Unicellular Cyanobacteria II Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. *Journal of Biological Chemistry* **266**:9515-9527

Page, R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**:357-358

Palenik, B. (1992) Polymerase evolution and organism evolution. *Current Opinion in Genetics and Development* **2**:931-936

Palenik, B. (1994) Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Applied and Environmental Microbiology* **60**:3212-3219

Palys, T., L.K. Nakamura and F.M. Cohan (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *International Journal of Systematic Bacteriology* **47**:1145-1156

Partensky, F., N. Hoepffner, W.K.W. Li, O. Ulloa and D. Vaulot (1993) Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. *Plant Physiol.* **101**:285-296

Rippka, R., et al. (2000) *Prochlorococcus marinus* Chisholm et al. 1992, subsp. nov. *pastoris*, strain PCC9511, the first axenic chlorophyll a_2/b_2-containing cyanobacterium (Oxyphotobacteria). *International Journal of Systematic Bacteriology* submitted

Rocap, G., L.R. Moore and S.W. Chishom (1999) Molecular Phylogeny of *Prochlorococcus* Ecotypes. In: Charpy, L. and A.W.D. Larkum (ed) Marine Cyanobacteria. Bulletin de l'Institut Océanographique, Monaco, special issue N° 19 (632 pages), p 107-116

Scanlan, D., W. Hess, F. Partensky, J. Newman and D. Vaulot (1996) High degree of genetic variation in *Prochlorococcus* (Prochlorophyta) revealed by RFLP analysis. *European Journal of Phycology* **31**:1-9

Shalapyonok, A., R.J. Olson and L.S. Shalapyonok (1998) Ultradian growth in *Prochlorococcus* spp. *Applied and Environmental Microbiology* **64**:1066-1069

Shimada, A., S. Kanai and T. Maruyama (1995a) Partial sequence of Ribulose-1,5-Bisphosphate Caboxylase/Oxygenase and the physiology of *Prochloron* and *Prochlorococcus* (Prochlorales). *Journal of Molecular Evolution* **40**:671-677

Shimada, A., M. Nishijima and T. Maruyama (1995b) Seasonal appearance of *Prochlorococcus* in Suruga Bay, Japan in 1992-1993. *Journal of Oceanography* **51**:289-300

Smith, S.W., R. Overbeek, C.R. Woese and W. Gilbert (1994) The genetic data environment: an expandable GUI for multiple sequence analysis. *CABIOS* **10**:671-675

Swofford, D.L. (1999) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts

Tay, S.T.-L., H.H. Hemond, L.R. Krumholz, C.M. Cavanaugh and M.F. Polz (2000) Population dynamics of two toluene degrading bacterial species in a contaminated stream assessed by quantitative PCR. submitted

Toledo, G. and B. Palenik (1997) *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. *Applied and Environmental Microbiology* **63**:4298-4303

Toledo, G., B. Palenik and B. Brahamsha (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. *Applied and Environmental Microbiology* **65**:5247-5251

Tomioka, N. and M. Sugiura (1984) Nucleotide Sequence of the 16S-23S Spacer region in the *rrnA* operon from a blue-green alga, *Anacystis nidulans*. *Molecular and General Genetics* **193**:427-430

Urbach, E. and S.W. Chisholm (1998) Genetic diversity in *Prochlorococcus* populations flow cytometrically sorted from the Sargasso Sea and the Gulf Stream. *Limnology and Oceanography* **43**:1615-1630

Urbach, E., D. Distel, D.J. Scanlan, J.B. Waterbury and S.W. Chisholm (1998) Rapid diversification of marine picophytoplankton with dissimilar light harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *Journal of Molecular Evolution* **46**:188-201

Urbach, E., D. Robertson and S.W. Chisholm (1992) Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature* **355**:267-270

Veldhuis, M.J.W. and G.W. Kraay (1993) Cell abundance and fluorescence of picophytoplankton in relation to growth irradiance and nitrogen availability in the Red Sea. *Neth. J. Sea. Res.* **31**:135-145

Veldhuis, M.J.W., G.W. Kraay, J.D.L. Van Bleijswijk and M.A. Baars (1997) Seasonal and spatial variation in phytoplankton biomass, productivity and growth in the northwestern Indian Ocean: the southwest and northeast monsoon, 1992-1993. *Deep Sea Research* **44**:425-449

Ward, D.M. (1998) A natural species concept for prokaryotes. *Current Opinion in Microbiology* **1**:271-277

Ward, D.M., M.J. Ferris, S.C. Nold and M.M. Bateson (1998) A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiology and Molecular Biology Reviews* **62**:1353-1370

Ward, D.M., R. Weller and M.M. Bateson (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**:63-65

Waterbury, J.B., S.W. Watson, F.W. Valois and D.G. Franks (1986) Biological and ecological characterization of the marine unicellular cyanobacteria *Synechococcus*. *Can. Bull. Fish. Aquat. Sci.* **214**:71-120

Waterbury, J.B., J.M. Willey, D.G. Franks, F.W. Valois and S.W. Watson (1985) A cyanobacterium capable of swimming motility. *Science* **30**:74-76

West, N.J. and D. Scanlan, J, (1999) Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Applied and Environmental Microbiology* **65**:2585-2591

Willey, J.M. and J.B. Waterbury (1989) Chemotaxis towards nitrogenous compounds by swimming strains of marine *Synechococcus* spp. *Applied and Environmental Microbiology* **55**:1888-1894

CHAPTER FIVE

Genetic diversity of *Prochlorococcus* and *Synechococcus* populations in the Sargasso Sea

ABSTRACT

The diversity of Sargasso Sea cyanobacterial populations was examined in summer and winter by restriction length polymorphism and sequencing analysis of the 16S/23S ribosomal DNA internal transcribed spacer (ITS). Cyanobacterial ITS sequences were amplified from total community genomic DNA in three different samples, collected to represent different temperature and nutrient regimes. PCR products were cloned and screened with two tetrameric restriction enzymes. A combined total of 152 clones from the three clone libraries yielded 19 different RFLP types. ITS sequences from representatives of each RFLP type were compared to sequences from *Prochlorococcus* and *Synechococcus* cultures and all but one clone could be assigned to one of seven ecotypic clusters. The single leftover clone is has a relatively low sequence similarity (~60%) to other environmental or cultured sequences of marine cyanobacteria. A larger number of *Prochlorococcus* clones were recovered compared to *Synechococcus* than would be expected from the relative cell abundances of each genus in the original field samples.

INTRODUCTION

Primary production in the open ocean is dominated by the unicellular cyanobacteria *Prochlorococcus* and *Synechococcus*. Although closely related phylogenetically (Palenik and Haselkorn 1992, Urbach, et al. 1992), they differ in their cell sizes and pigment complements, enabling them to be easily distinguished and enumerated by flow cytometry (Olson, et al. 1990a, Olson, et al. 1990b). Together *Prochlorococcus* and

Synechococcus comprise up to 50% of carbon fixation in oligotrophic waters (Campbell and Vaulot 1993, Goericke and Repeta 1993, Veldhuis, et al. 1993, Binder, et al. 1996). Isolates of *Prochlorococcus* and *Synechococcus* with different photosynthetic capacities have been shown to coexist in natural samples (Toledo and Palenik 1997, Moore, et al. 1998). Knowledge of the genetic diversity, and corresponding physiological diversity, of these important photoautotrophs will provide a better understanding of the processes that influence rates of photosynthesis in the ocean.

Prochlorococcus isolates can be divided into two physiologically and genetically distinct groups, referred to as ecotypes because their differing physiologies have implications for their ecological distributions. The primary differentiator of the ecotypes is their chl b/a_2 ratio, although they also differ in their light-dependent physiological responses (growth rate, pigment content, and photosynthetic rate as a function of light intensity). Isolates with a high chl b/a_2 ratio (high B/A ecotype) are much more efficient at utilizing low light than those with a low chl b/a_2 ratio (low B/A ecotype), but are incapable of growth at higher irradiances (Moore and Chisholm 1999). In addition, high B/A isolates are much more sensitive to copper than isolates of the low B/A ecotype (Mann 2000). Analysis of 16S rDNA and 16S 23S internal transcribed spacer (ITS) sequences demonstrates that low B/A isolates are closely related and form a well supported clade, which can be further divided into two clusters (Urbach, et al. 1998, Rocap, et al. 1999). In contrast, the high B/A isolates are divided among of four clusters which are less

similar in their ITS sequence and do not form a monophyletic clade (Rocap, et al. 1999, Chapter 4).

Marine *Synechococcus* strains are also physiologically and genetically diverse. *Synechococcus* can be distinguished based on the pigment composition of their light harvesting complexes, as isolates vary in the amounts of the chromophores phycourobilin (PUB) and phycoerythrobilin (PEB) which attach to the phycobiliprotein phyocerythrin (Waterbury, et al. 1986). Strains also exhibit other physiological differences in swimming motility, nutrient utilization and cell cycle behavior (Waterbury, et al. 1986, Binder and Chisholm 1995). Investigations into the genetic diversity of *Synechococcus* using DNA-dependent RNA polymerase gene sequences (*rpoC1*) as well as the ITS have identified numerous strain clusters (Toledo and Palenik, 1997, Toledo et al. 1999, Chapter 4). However the relationship of these strain clusters to physiological diversity is not completely clear, except in the case of one cluster which contains all the motile isolates examined to date (Toledo and Palenik, 1997, Toledo et al. 1999, Chapter 4).

A major issue in extrapolating physiological properties of cultured isolates to natural populations is whether the strains represent dominant members of environmental communities. Sequence diversity has been assessed in field populations of marine cyanobacteria using *rpoC1* sequences (Palenik 1994, Ferris and Palenik 1998) and in *Prochlorococcus* specifically using 16S rDNA (West and Scanlan 1999) and the intergenic region between the photosynthetic electron transport genes *petB* and *petD*

(*petB/D*) (Urbach and Chisholm 1998). PCR amplified 16S rDNA sequences from a North Atlantic water column could all be affiliated with two *Prochlorococcus* clusters (high B/A I and low B/A I) (West and Scanlan 1999). However, the extent of physiological diversity represented by the genetic diversity found at *rpoC1* or *petB/D* is not clear, partly because of the relatively few *Prochlorococcus* isolates sequenced at these loci, and also because the connection between physiological and genetic diversity in *Synechococcus* is still being established (Toledo and Palenik 1997, Toledo, et al. 1999).

The internal transcribed spacer (ITS) region between the 16S and 23S rRNAs has recently been used to describe the multiple ecotypes of *Prochlorococcus* and *Synechococcus* (Chapter 4). The higher variability of this region compared to the 16S rDNA allowed the resolution of ten sequence similarity clusters, each of which likely corresponds to an ecologically distinct population. The aims of this study were to characterize the sequence diversity of natural cyanobacterial populations at the ITS, using the sequences from *Prochlorococcus* and *Synechococcus* cultures as a framework to attach a phenotype to environmental genotypes.

MATERIALS AND METHODS

Sampling locations. Water samples were collected on two cruises to the Sargasso Sea aboard the RV Oceanus. PCR clone libraries were constructed from samples obtained from 32° 08' N, 70°02' W at 75 m on June 10, 1996 (OC279 station 6 CTD 31) and from 35°48' N, 61°56' W at 15 m and 150 m on January 30, 1997 (OC297 station 5 CTD 12).

Water was collected from Niskin bottles. Temperature and in vivo chlorophyll fluorescence data were recorded by a SeaBird CTD mounted to the sampling rosette.

Nutrient data presented are from the same cruises and stations but different CTDs. June data is from 32° 18' N, 70°11' W (OC279 station 6 CTD 18) and January data was obtained from 35°.83 N and 61°.95 W (OC297 station 5 CTD 31). The collection and measurement of inorganic nitrogen (nitrate and nitrite, or N+N) and soluble reactive phosphorous (SRP) is described by Cavender-Bares (1999).

Cell counts. Seawater (1.2 ml) was preserved for flow cytometry in 0.125% glutaraldehyde (final concentration) and stored in liquid nitrogen. Samples were analyzed on a modified FACScan flow cytometer (Dusenberry and Frankel 1994) using Cyclops software (Cicero).

DNA isolation. Immediately after collection between 1-2 l of seawater was filtered on to a 0.2 μm filter (nucleopore) and stored in liquid nitrogen. For DNA extraction, filters were thawed, sliced in half, and one half was refrozen. The half used for extraction was minced and added to 2 ml lysis solution (0.5 M NaCl, 10 mM EDTA, 10 mM Tris pH 8.0), and 0.2 ml 10% SDS was added dropwise while vortexing. The sample was boiled for 2 minutes, spun 10 min in a microcentrifuge at 10k rpm, and the supernatant was transferred to new tubes. The sample was extracted twice with 25:24:1 phenol: chloroform: isoamyl alcohol and once with 24:1 chloroform: isoamyl alcohol, and then

0.12 ml 10 M ammonium acetate and 0.75 ml ice cold ethanol were added. DNA was precipitated at –20°C overnight and then spun 10 min in a microfuge at 4°C, washed with 70% ethanol, dried at room temperature and resuspended in TE.

PCR clone library construction. Portions of the 16S and 23S rDNAs and the entire ITS between them (the ITS/23S locus) were amplified from the three samples using primers 16S-1247f (CGTACTACAATGCTACGG) and 23S-1608r (CYACCTGTGTCGGTTT). Primer 16S-1247f is cyanobacterial specific, as determined using the PROBE-CHECK function in the newest release of the RDP (Maidak, et al. 1999). PCR reactions employed cycling parameters described in Chapter 4. Reactions were performed in quintuplicate, pooled, and excess primers removed using Strataprep columns (Stratagene). An aliquot was visualized on a 1% agarose gel to ensure that only bands in the 2300-2700 bp size range were present. Control reactions lacking template DNA were always performed in parallel and gave no products, as judged by visualization on agarose gels.

Pooled purified PCR products were cloned into pCR-Script vectors (Stratagene) and transformed into Super competent cells (Stratagene) according to the manufacturer's instructions. Positive colonies, as judged by blue-white color selection, were picked and struck on selective plates to confirm the color phenotype and obtain single colonies, which were picked and grown overnight in LB supplemented with appropriate antibiotics. These liquid cultures were used for frozen stocks (mixed 1:1 with glycerol storage buffer and stored at -80°C (Ausubel, et al. 1992)) and for plasmid purification using the

Miniprep–24 (MacConnell Research). Plasmids were digested with Eco RV (New England Biolabs), which cuts the pCR-Script plasmid once (in the polylinker). Plasmids with appropriate sized inserts (roughly 50%) were selected for further analysis. The majority of inserts were not cut by Eco RV.

Restriction fragment length polymorphism (RFLP) analysis. Minipreps (0.1-1 µl) were used as templates to amplify the ITS/23S from each plasmid using the PCR conditions and primers described above, except only single reactions were performed for each template. PCR products (3-8 µl) were independently digested in 20 µl reactions with the tetrameric restriction enzymes Hae III (GG/CC) and Rsa I (GT/AC) (New England Biolabs) according to manufacturer's instructions. Digests were visualized on 2% agarose TBE gels stained with ethidium bromide. Fragments were sized using Eagle Eye software (Stratagene) and only those larger than 200bp were considered in resolving the RFLP patterns. The combined results of the two single digests were used to determine RFLP type.

For RFLP analysis of cultures, isolates of *Prochlorococcus* and *Synechococcus* were grown and genomic DNA was extracted as described in Chapter 4. Amplification of the ITS/23S and determination of the RFLP type for each culture was performed as above.

Sequencing. A 1000-1300 bp fragment spanning the ITS was sequenced in at least two clones of each RFLP type (where possible) from each sample (51 clones total). PCR

products from each plasmid obtained as above were sequenced on one strand using primers 16S-1247f and 23S-241r (TTCGCTCGCCRCTACT). Sequencing reactions were performed with ABI Big Dye terminator sequencing kits (PE Biosystems) according to the manufacturer's instructions and run on an ABI310 automated sequencer. Sequences of five clones from the Jan. 150 m clone library (47, 59, 144, 145, 150) whose phylogenetic affiliation was initially uncertain were confirmed by second strand sequencing using primers ITS-alaf (TWTAGCTCAGTTGGTAGAG) and ITS-alar (CTCTACCAACTGAGCTAWA). Base calls in sequence chromatagram files were manually checked using Chromas v. 1.43 (publicly available at http://www.technelysium.com.au/chromas.html).

Phylogenetic analysis. Sequences were compiled and aligned manually with ITS sequences of cultured cyanobacteria (Chapter 4) using the Genetic Data Environment (Smith, et al. 1994). Phylogenetic analyses employed PAUP* version 4b2a (Swofford 1999). Distance trees were inferred using minimum evolution as the objective criterion and paralinear (logdet) or HKY85 distances. Maximum likelihood analyses used the HKY85 model of nucleotide substitution with rate heterogeneity and empirical nucleotide frequencies. The gamma shape parameter and the transition/transversion ratio were initially estimated from a distance topology and refined by iterative likelihood searches. All heuristic searches, including bootstrap analyses (100 resamplings), utilized random addition and tree-bisection reconnection (TBR) branch-swapping methods. Phylogenetic trees were visualized with TREEVIEW (Page 1996).

RESULTS

Sampling locations. Seawater samples were collected on two cruises to the Sargasso Sea in June 1996 and January 1997. The water column in June exhibited a typical summer profile for the Sargasso Sea (Olson, et al. 1990b) with a shallow mixed layer and a deep chlorophyll maximum (Fig. 1A). Nutrient concentrations were low in surface waters (Fig. 1B) (Cavender-Bares 1999). Concentration of *Synechococcus* were fairly uniform throughout the water column while *Prochlorococcus* were depleted in surface waters and exhibited a subsurface maximum (Fig. 1C).

In contrast, the water column in January exhibited a deep mixed layer, typical for winter in the Sargasso Sea (Fig. 1D) (Olson, et al. 1990b). Consistent with the constant temperature over the upper 150 m, nutrient concentrations were relatively constant with depth. Nitrogen concentrations in surface waters were 100 fold greater than in the June profile (Fig. 1E) (Cavender-Bares 1999). Cell numbers for both *Prochlorococcus* and *Synechococcus* were also relatively constant with depth. *Prochlorococcus* were ~20% more abundant than *Synechococcus* throughout water column (Fig. 1F).

PCR clone library construction. Three samples from the water columns described above were used to construct PCR clone libraries (Fig. 1C, F). In June a sample from 75 m, where *Prochlorococcus* had a maximum in cell numbers, was successfully used to create a PCR clone library. Samples were also taken from 15 m and 25 m, but due to extremely low DNA recovery from these samples, cyanobacterial ITS sequences could

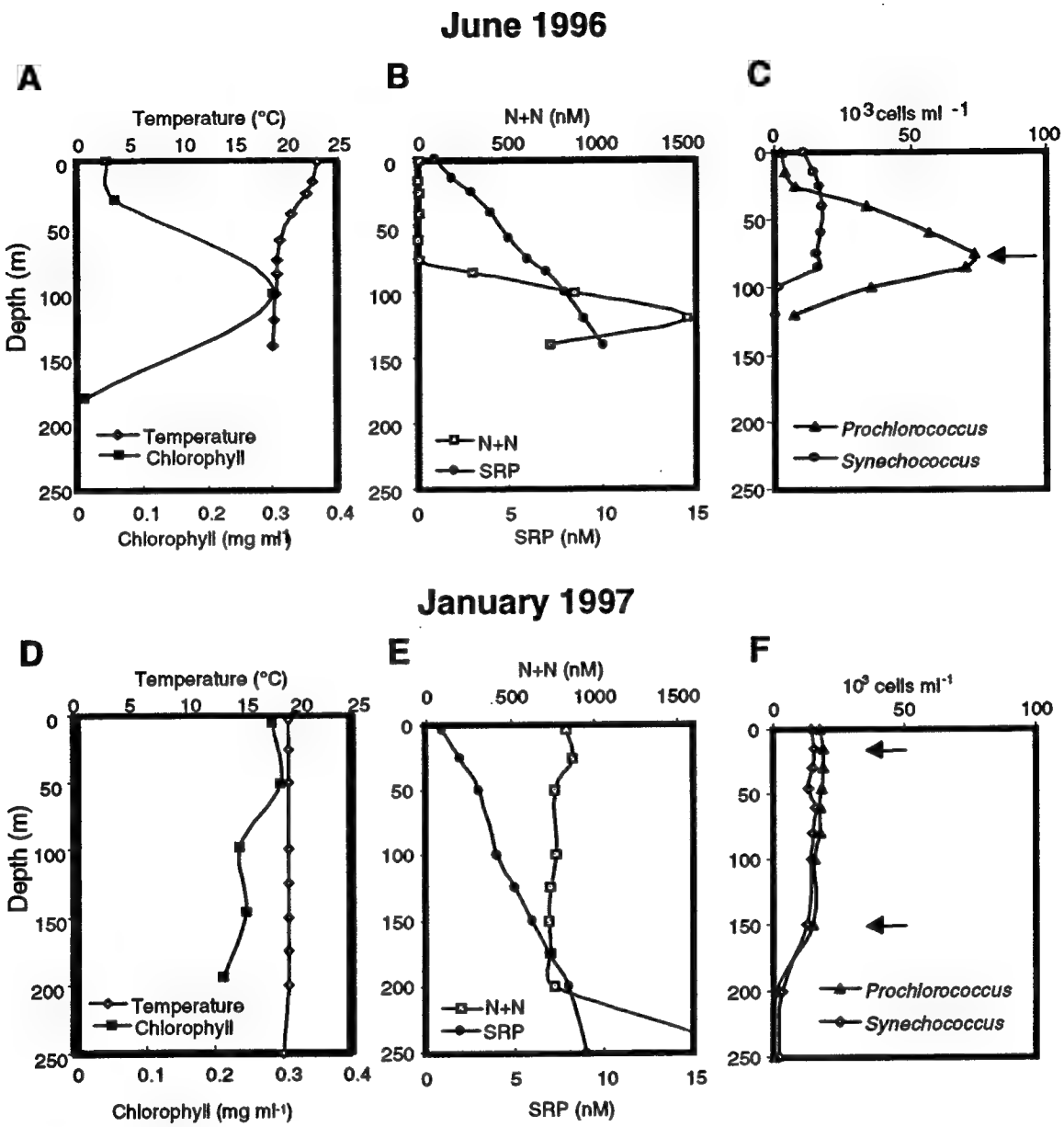


Figure 1. Water column characteristics at sampling locations from summer (A, B, C) and winter (D, E, F) cruises. **(A, D)** Temperature and chlorophyll with depth. **(B, E)** Total nitrogen (nitrate and nitrite (N+N)) and soluble reactive phosphorus (SRP) (from Cavender-Bares, 1999)). **(C, F)** Total abundance of *Prochlorococcus* and *Synechococcus*. Arrows indicate sampling locations for PCR clone libraries.

not be reproducibly amplified, and no PCR clone libraries were constructed from surface waters of this profile. In January samples from near the surface (15 m) and at the base of the mixed layer (150 m) were employed as material for PCR clone library construction.

The ITS/23S locus was amplified from nucleic acids extracted from the three samples using primers to the 16S and 23S rDNAs. Only the forward (16S rDNA) primer was cyanobacterial specific, but no sequences suspected to originate from heterotrophic bacteria were found in any of the libraries. The PCR products were cloned and 152 clones (48, 32 and 72 from the June 75 m, Jan. 15 m and Jan. 150 m libraries respectively) were recovered and subjected to further analysis.

Determination of RFLP types. The tetrameric restriction enzymes Hae III and Rsa I were originally chosen to screen the environmental sequences because restriction patterns of ITS/23S sequences distinguish high B/A *Prochlorococcus* from low B/A isolates (Fig. 2). The enzymes are also able to resolve the clusters within the *Prochlorococcus* ecotypes (Fig. 2), and clusters of *Synechococcus* isolates (data not shown).

Amplified ITS sequences from clones in the three environmental libraries were screened independently with Hae III and Rsa I. In several cases clones were found with identical restriction patterns with one of the enzymes and different banding patterns with the other enzyme; these were designated separate RFLP types. For example, types C and D show the same restriction pattern with Rsa I (Fig. 3A) but a different one when Hae III is used

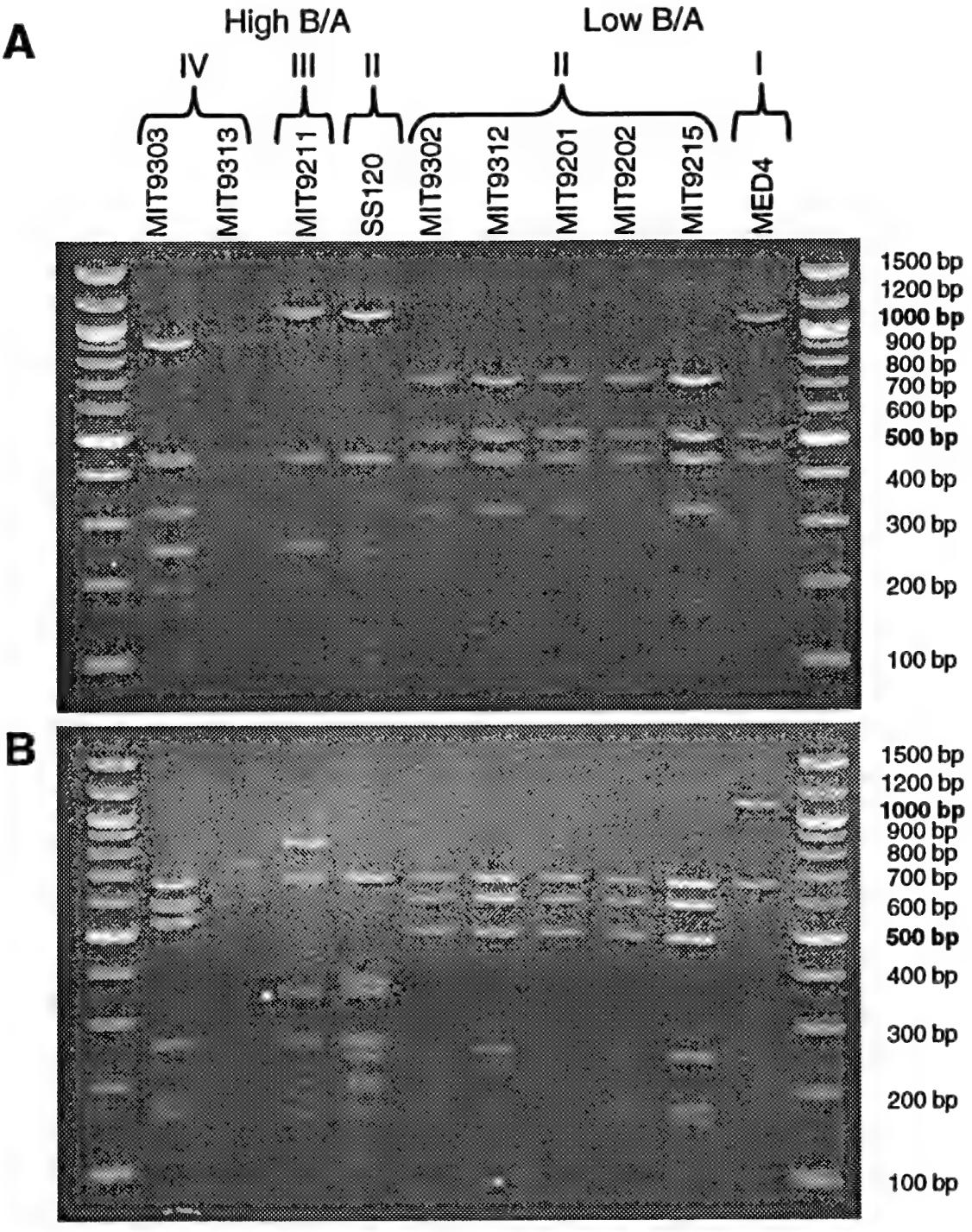


Figure 2. Restriction digests of cultured *Prochlorococcus* ITS/23S sequences. **(A)** Digests using Rsa I. **(B)** Digests using Hae III.

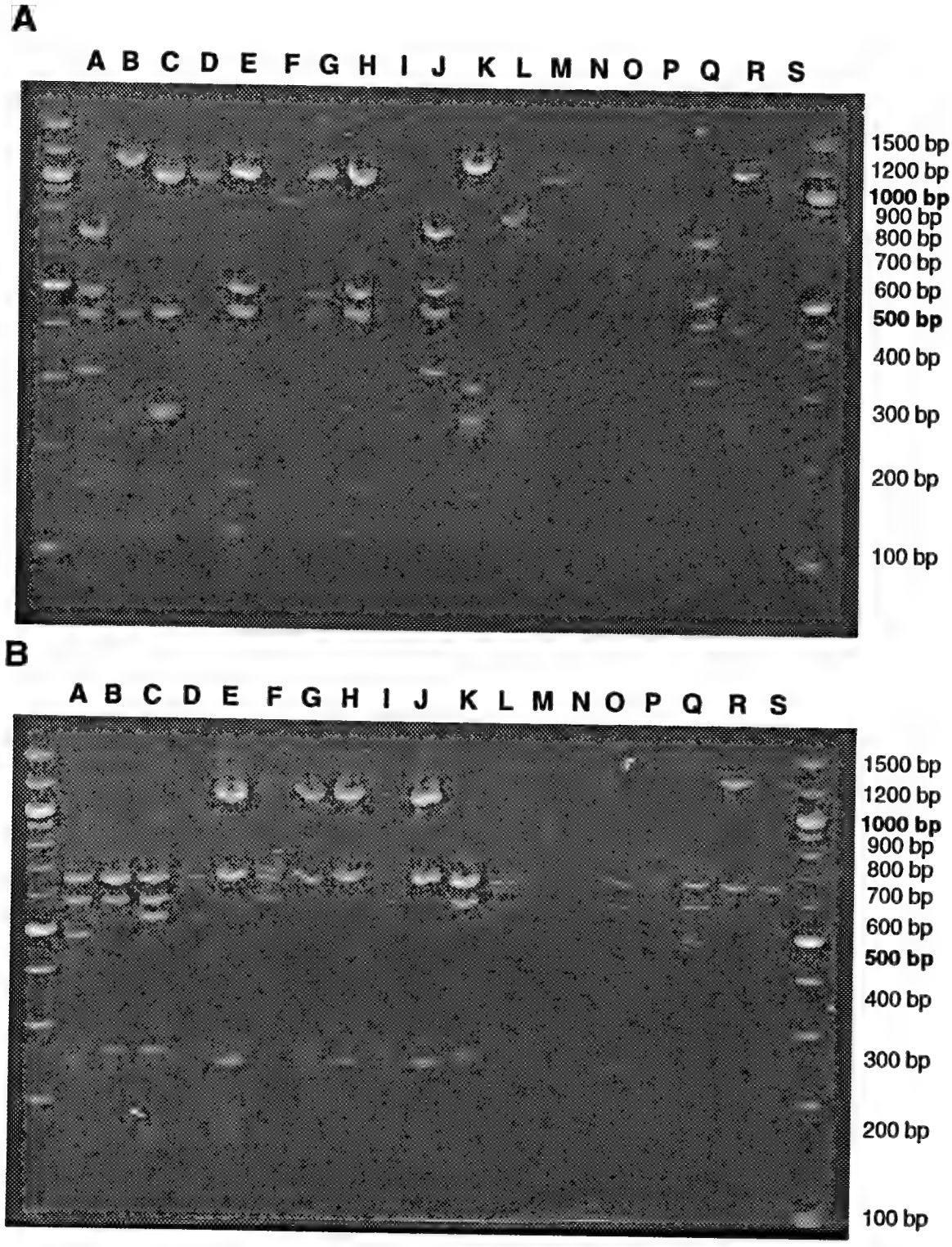


Figure 3. Restriction digests of the ITS/23S in environmental clones representative of each RFLP type found. (**A**) Rsa I digests (**B**) Hae III digests

(Fig. 3B). Based on the combined patterns of the two single digests a total of 19 RFLP types were identified among the three samples (Fig. 3). The June 75 m library contained 9 RFLP types, the Jan. 15 m library yielded 4 types and clones in the Jan. 150 m library consisted of 12 types (Table 1).

Diversity and distribution of RFLP types. Of the 19 RFLP types recovered in the three libraries, 2 RFLP types (A and B) were found in all three libraries and 2 (C and E) were found in two of the three libraries (both in the June 75 m and the Jan. 150 m libraries) (Table 1). The distribution of clones among the RFLP types was quite skewed, with the 4 RFLP types found in more than one library accounting for 126 of the 152 clones examined. The remaining 15 RFLP types were each represented by 4 clones or fewer, with 9 RFLP types found only once (Table 1).

Seven of the nineteen environmental RFLP types were also found in cultures. The four most prevalent types (A, B, C and E) all had counterparts in cultures (*Prochlorococcus* MIT9312, *Synechococcus* WH8012, *Prochlorococcus* NATL2A and *Prochlorococcus* MED4, respectively) (Table 1). Patterns identical to *Prochlorococcus* NATL1A, *Synechococcus* WH8103 and *Synechococcus* WH9908 were also found in the clone library (RFLP types D, F and S) (Table 1).

Table 1. Clones assigned to each RFLP type. For each field sample library clones were numbered sequentially in the order they were picked off solid media. Numbers shown in bold were selected for sequencing.

RFLP type	June 75 m	Jan 15 m	Jan 150 m	number sequenced
A	**3**,6,**10**,13,**15** 16,17,29,33, 35,51,60,62, 66	**26**,29,48,**52**,56, 65,67,71,88,90, 108,112,**127**, 137,167,171, 177	**4**,20,21,26,27,**33**,34 39,41,48,52,62,63, **66**,68,73,77,79,82, 87,90,96,**100**,104, 111,118,123,125,12 7,131,132,143,146, 148,149,152,**155**	11
B	4, **26**	**3**,**47**,59,93,109, 123,**124**,125, 134,138,140, 141,**156**	**15**,17,19,28,42,46, **49**,**101**,**103**,107,109 110,120,121,134, **141**,142,151,**153**	11
C	1,5,**7**,**18**,19, 30,31,43,44, 47, 54,58		**13**,**106**,**124**	5
D	**37**,**45**,**46**, 61			3
E	34,**49**,**53**,55, 56,57, 59,64		**1**	3
F	27,**32**			1
G	**14**,28,38			1
H	20,**42**			1
I	**63**			1
J		**96**		1
K		**187**		1
L			**59**,**144**,**145**	3
M			**86**,**88**,**112**	3
N			**47**	1
O			**10**	1
P			**56**	1
Q			**78**	1
R			**115**	1
S			**150**	1
Total				51

In order to assess how well the clones screened represented the diversity present in the clone libraries the cumulative numbers of RFLP types were plotted as they appeared during sampling (Fig. 4) (Moyer, et al. 1994). As a library becomes more thoroughly sampled, the examination of more clones will not result in novel RFLP types recovered and the cumulative plots will reach a plateau corresponding to the actual number of types in the library. Since the curves should asymptotically approach the true number of RFLP types in the library their shape is analogous to Michalis-Menten curve and therefore in a double-reciprocal plot, analogous to a Lineweaver-Burke plot, the reciprocal of the y-intercept is the total number of RFLP types (Polz, et al. 1999). Double reciprocal plots of the data in Fig. 2 yield y-intercept values of 8.4, 3.0 and 11.6 for the June 75 m, Jan. 15 m and Jan. 150 m libraries respectively. These are quite similar to the observed values (9, 4 and 12) indicating that the libraries were almost completely sampled. It should be emphasized that, because of the biases that can occur in steps used to construct the library (chiefly the PCR), this coverage refers to the completeness of sampling the diversity of clones in the library, but cannot be extended to conclude that the sequence diversity in the environment has been fully examined.

Phylogenetic Analysis. To confirm the close relationship of seven of the RFLP types to their respective cultures and to assign a phenotype to the remaining twelve RFLP patterns, the ITS was sequenced in 51 clones (Table 1). ITS sequences varied in length from 539 to 811 bp, within the range of ITS lengths in cultures of *Prochlorococcus* and *Synechococcus* (Chapter 4). Eleven clones were sequenced of each of the two most

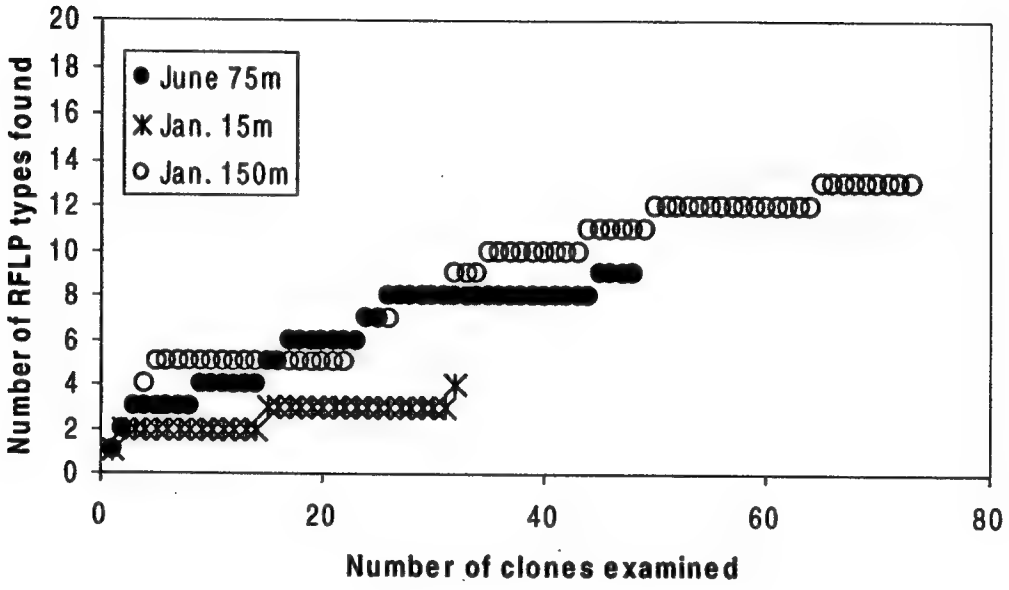


Figure 4. Estimated diversity of cyanobacterial RFLP types in three clone libraries from the Sargasso Sea. For each library the cumulative number of RFLP types found is graphed as a function of the number of clones screened.

abundant RFLP types (Table 1). Within a single RFLP type sequence similarities at the 16S-tRNA^{Ile} spacer ranged from 93-100%.

Phylogenetic trees based on 103 positions of the 16S-tRNA^{Ile} spacer show that the majority of the clones fall into previously defined clusters based on sequences from cultures (Fig. 5). The nineteen restriction patterns were grouped in ecotypes based on sequence similarities and phylogenetic position of the sequences (Table 2). Seven of the eight clusters have cultured members. In many cases these groups were anticipated even before sequencing because clones had identical RFLP patterns with one restriction enzyme (Fig. 3).

Twelve of the RFLP types can be identified as *Prochlorococcus*. Nine of these branch within the low B/A *Prochlorococcus* clade. Three of the low B/A RFLP types (E, G, and H) are members of cluster I and the other six (A, J, M, O, P, Q) belong to the low B/A II cluster (Table 2, Fig. 5). High B/A *Prochlorococcus* are represented in the libraries by three RFLP types. Two of these (C and D) share RFLP patterns with the high B/A I cluster cultures NATL2A and NATL1A, respectively (Table 2). A third RFLP type (I) also branches within the high B/A cluster I (Fig. 5). However, no clones were found with RFLP patterns like isolates of the other three high B/A *Prochlorococcus* clusters and none of the sequenced clones had high sequence similarities to those clusters.

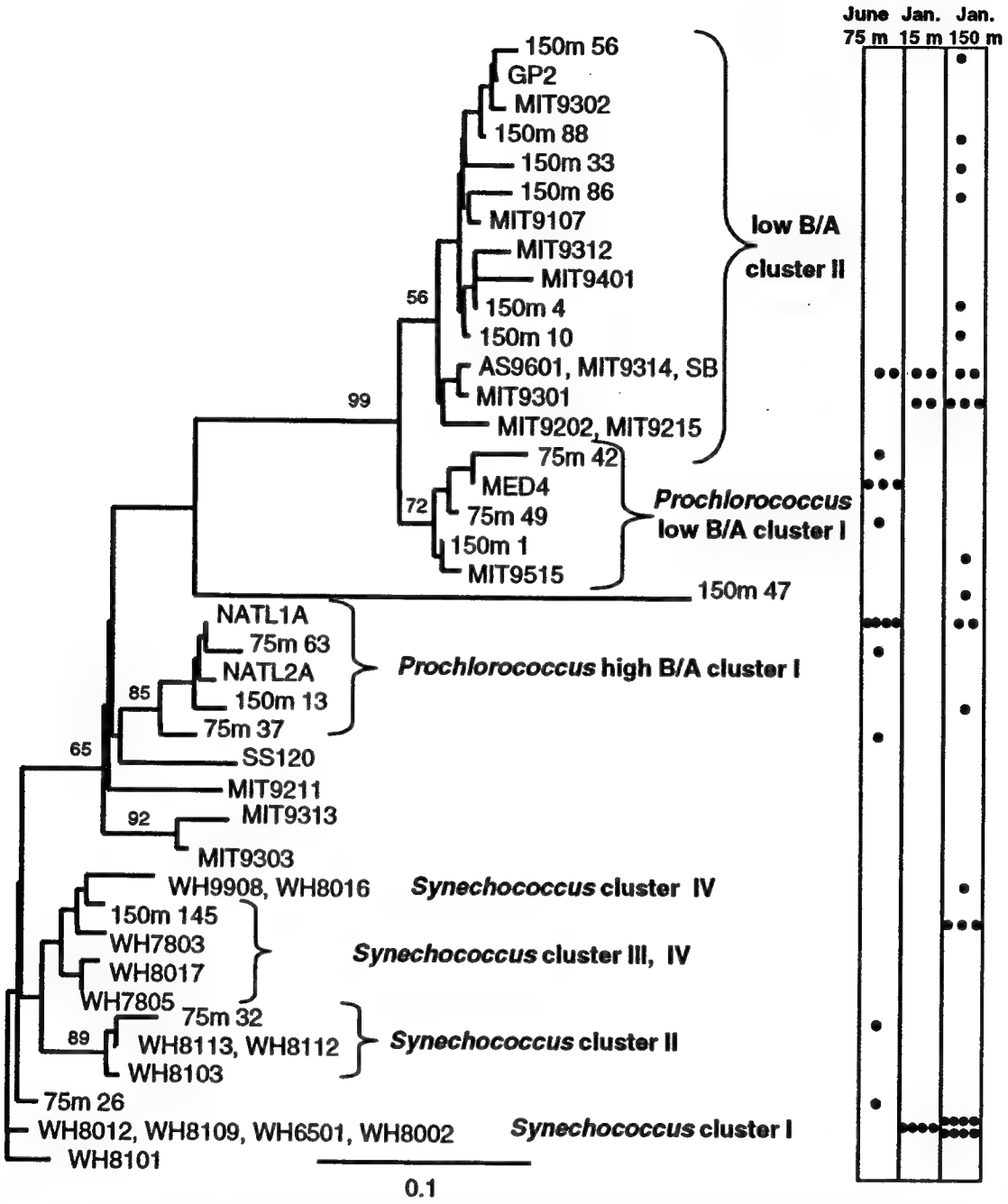


Figure 5. Evolutionary relationships of 51 environmental sequences with cultured isolates of marine cyanobacteria. Dots indicate sequences from clone libraries on tree, more than one dot on a line indicates sequences identical at the 103 positions of the16S-tRNA^{Ile} spacer used in the analysis. Phylogenetic framework inferred using a distance criterion with paralinear (logdet) distances and minimum evolution as the objective function. Bootstrap values from 100 resamplings are to the left of each nod, values below 50 not shown. The tree found using likelihood methods assigned clones to same sequence clusters and was not significantly better as judged by the Kishino and Hasegawa test. Trees excluding low B/A *Prochlorococcus* and employing 146 positions of the 16S-tRNA^{Ile} spacer resulted in essentially similar branching orders and exactly the same cluster assignments.

Table 2. Distribution of RFLP types among ecotypes

Ecotypic cluster	Representative culture	RFLP type	Numbers of clones found: June 75m	Jan. 15m	Jan. 150m	Total
Prochlorococcus Low B/A II	MIT9312	A	14	17	37	68
		J		1		1
		M			3	3
		O			1	1
		P			1	1
		Q			1	1
Prochlorococcus Low B/A I	MED4	E	8		1	9
		G	3			3
		H	2			2
Prochlorococcus High B/A I	NATL2A	C	12		3	15
	NATL1A	D	4			4
		I	1			1
Synechococcus I	WH8012	B	2	13	19	34
		K		1		1
		R			1	1
Synechococcus II	WH8103	F	2			2
Synechococcus III		L			3	3
Synechococcus IV	WH9908	S			1	1
Unaffiliated		N			1	1
Total			48	32	72	152

Six of the nineteen environmental RFLP types could be designated as *Synechococcus* (Table 2, Fig. 5). Three different RFLP types (B, K, and R) branched within *Synechococcus* cluster I, which consists of strains with low PUB levels (Fig. 5). A fourth RFLP type (F) is shared by the motile *Synechococcus* in cluster II. A fifth RFLP type (L)

branched within *Synechococcus* cluster III. Finally the sixth RFLP type (S) was identical to that of isolates in *Synechococcus* cluster IV, and sequence similarity bore this out.

The nineteenth RFLP type (N) could not be affiliated with any cultures. Clone number 47 from the Jan. 150 m sample is only 50-60% similar at the 16S-tRNA^{Ile} spacer to high B/A *Prochlorococcus* and *Synechococcus* (as compared to more than 90% similarity within clusters and generally greater than 75% similarity between clusters). Although phylogenetic analyses place this sequence within the high B/A *Prochlorococcus*, this branching pattern is not well supported. The unstable position of this clone is likely due to its large genetic distance from all of the other sequences (Swofford, et al. 1996).

The BLAST searching function of GenBank was used to further explore the possible identity of cloned sequence 150 m-47. A search using the full ITS sequence resulted in a collection of cyanobacterial and chloroplast sequences. However, these same sequences were also identified as the best matches if any of the cultured *Prochlorococcus* or *Synechococcus* sequences were utilized as input data to the search. Thus clone 150 m-47 is likely cyanobacterial in origin, but the specific hits reflect the limited number of cyanobacterial ITS sequences in the database rather than a particular affiliation for this clone. If the search was restricted to the approximately 150 bp on the 3' end of the 16S rDNA, then the best match was the 16S rDNA sequence from the high B/A *Prochlorococcus* isolate SS120, followed by other *Prochlorococcus* and marine *Synechococcus* isolates.

Biases in distribution of clones. Because of their distinctive pigment fluorescences *Prochlorococcus* and *Synechococcus* populations can be enumerated individually by flow cytometry, thus it is possible to compare the relative proportions of clones assigned to each of the two genera with the proportions of cells in the sample each library was constructed from. This calculation provides some indication of bias in the clone libraries. All three libraries over-represent *Prochlorococcus* with respect to *Synechococcus* by factors ranging from 1.1 to 2.3 (Table 3).

Table 3. Estimates of relative bias in clone libraries. *Prochlorococcus* overestimation factor was determined for each sample by dividing the ratio of *Prochlorococcus* to *Synechococcus* (Pro:Syn) in the PCR clone libraries by the ratio of Pro:Syn cell numbers.

	June 75m	Jan. 15 m	Jan. 150 m*
PCR clone libraries			
Prochlorococcus	44	18	48
Synechococcus	4	14	23
Pro:Syn ratio	**11 : 1**	**1.3 : 1**	**2.09 : 1**
Cell concentrations (ml^{-1})			
Prochlorococcus	73843	19000	15400
Synechococcus	15700	15900	13100
Pro:Syn ratio	**4.7 : 1**	**1.19 : 1**	**1.17 : 1**
Prochlorococcus overestimation factor	2.34	1.1	1.79

*Only 71 clones are included in analysis of Jan. 150 m clone library as one clone could not be confidently assigned to either *Prochlorococcus* or *Synechococcus*

DISCUSSION

The coexistence of multiple cyanobacterial ecotypes in single water samples has been repeatedly demonstrated with both culture and sequence data (Toledo and Palenik 1997, Ferris and Palenik 1998, Moore, et al. 1998, Urbach and Chisholm 1998). Here, of eight clusters found, two were found in all three libraries and two more were found in two of the three. These results are consistent with analysis of *petB/D* intergenic region sequences in the Sargasso Sea and Gulf Stream in which overlapping alleles were found in multiple libraries from a depth profile (Urbach and Chisholm 1998). This was interpreted as populations in the water column sharing a common gene pool, with different alleles potentially changing in numeric abundance as conditions changed. A similar situation may exist over the course of a year. It is likely that several ecotypes are present all year round, and thus are available to be detected using PCR-based methods. However, the relative abundances may change throughout the year as conditions change in the water column.

The majority of sequences in this study were related to an isolate in culture (more than 90% similar at the 16S-tRNA^{Ile} spacer). The high proportion of sequences related to isolates raises concerns about the possibility that the environmental PCR samples could have been contaminated with culture DNA (especially already amplified ITS sequences) existing in the lab. However, although a few environmental sequences are identical to those of cultures, the majority are merely quite similar, which would not be expected if they were contaminants. In theory these small differences could be the result of

microheterogeneity at multiple copies of the ITS in a single organism, which may not have been observed in sequences obtained from PCR products in Chapter 4. However, preliminary data from the complete genome of *Prochlorococcus* MED4 (http://bbrp.llnl.gov/jgi/microbial/prochlorococcus_homepage.html) and from southern blotting of a suite of isolates (Urbach, et al. 1998, West and Scanlan 1999) indicates that *Prochlorococcus* has only one ribosomal RNA operon. Further, contamination as a source can be absolutely ruled out for all of the *Synechococcus* sequences found, as all three clone libraries were constructed before any of the *Synechococcus* strains described in Chapter 4 were present in the lab.

One clone (150 m-47) was not at all closely related to any available sequences from cultures. Cyanobacterial sequences of uncertain phylogenetic affiliation have been observed previously in environmental clone libraries. The branching pattern of SAR7, a 16S rDNA sequence from the Sargasso Sea (Giovannoni, et al. 1990), cannot be confidently determined within the marine picoplankton clade (Urbach, et al. 1998, Rocap, et al. 1999). However, SAR7 is very likely either a high B/A *Prochlorococcus* or a *Synechococcus*, while clone 150 m-47 is different enough from both *Prochlorococcus* and *Synechococcus* sequences to suggest it may represent a different genus altogether. Although *Prochlorococcus* and *Synechococcus* are by far the numerically dominant cyanobacteria in the open ocean, they are not the only genera present. In the Pacific larger (~2-3 μm) orange fluorescing cells have been observed and tentatively identified as *Synechocystis* (Campbell, et al. 1997, Neveux, et al. 1999). It is certainly possible for the

ITS from such an organism to have been amplified in this study, as the primer 16S-1247f was designed to exclude heterotrophic bacterial sequences, but is not specific to the *Prochlorococcus*/marine *Synechococcus* clade.

The overrepresentation of *Prochlorococcus* relative to *Synechococcus* in the clone numbers as compared to the cell numbers is evidence that biases existed somewhere in the process of turning cells into PCR clone libraries. Isolation of DNA is one step where biases may have occurred, as *Prochlorococcus* and *Synechococcus* may have different DNA extraction efficiencies for the protocol employed. The PCR is another likely candidate for introducing bias (Suzuki and Giovannoni 1996, Polz and Cavanaugh 1998), although attempts were made to avoid conditions known to induce PCR bias where possible. Multiple PCR reactions were pooled before cloning to reduce the influence of early random events (Polz and Cavanaugh 1998). The number of cycles was limited to the amount (30) where PCR products were barely visible on agarose gels. However, other strategies such as starting with high template concentrations in reactions and avoiding degenerate primers (Polz and Cavanaugh 1998) were not feasible because of low recovery of DNA from environmental samples and the limited database of cyanobacterial 23S rDNA sequences available. The template of interest may also be an issue, as the differences in ITS length and %GC content between the two genera could lead to differential amplification, specifically preferential amplification of shorter templates, such as low B/A *Prochlorococcus* (Mutter and Boynton 1995). This appears to be a likely explanation for the results observed here, although multiple factors may

have contributed to the overrepresentation of *Prochlorococcus* in the PCR clone libraries. Similar biases undoubtedly exist within each genus (i.e. among the different ecotypes of *Prochlorococcus* and *Synechococcus*).

Because of the demonstrated and potential biases, information on the relative distribution of the ecotypes must be garnered using quantitative methods such as probing (Field, et al. 1997) or Q-PCR (Lee, et al. 1996, Tay, et al. 2000). The discrete number of clusters identified by the ITS will be advantageous for the design of ecotype specific oligonucleotides to quantitate abundances (see Chapter 6). The ultimate test of a sequence similarity cluster as an ecologically discrete population is if it has a unique distribution in time and space (Palys, et al. 1997). Thus, the ability to measure the abundances of each cluster could confirm that they represent ecologically distinct organisms as well as provide insights into factors that may regulate their distributions.

CONCLUSIONS

This study demonstrates the utility of the ITS for species level diversity studies, as environmental ITS sequences could be readily grouped into sequence clusters. Consistent with previous studies demonstrating that it is not possible to make quantitative conclusions about genotype distributions on the basis of frequencies in clone libraries, the number of clones found was skewed towards *Prochlorococcus*. However, despite the presumably different set of biases associated with culturing environmental microorganisms, the overwhelming majority of sequences found were closely related to

those of cultures. Confirmation of this result and exploration of the potential different distributions of the ecotypes awaits the development of quantitative methods for determining the abundances of the ecotypes of *Prochlorococcus* and *Synechococcus*.

ACKNOWLEDGEMENTS

This work was supported by NASA and the NSF. I thank M. Polz for access to an automated sequencer, generous use of other laboratory equipment and helpful discussion, W. Pope, E. Cohen and N. Herschenhous for help in clone library screening, N. Ahlgren for sequencing assistance, K. Cavender-Bares for nutrient data, and the captain and crew of the R/V Oceanus.

REFERENCES

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl (1992) Short Protocols in Molecular Biology. John Wiley & Sons, New York

Binder, B.J. and S.W. Chisholm (1995) Cell Cycle Regulation in Marine *Synechococcus* sp. Strains. *Applied and Environmental Microbiology* **61**:708-717

Binder, B.J., S.W. Chisholm, R.J. Olson, S.L. Frankel and A.Z. Worden (1996) Dynamics of Pico-phytoplankton, Ultra-Phytoplankton, and Bacteria in the Central Equatorial Pacific. *Deep Sea Research* **43**:907-931

Campbell, L., H. Liu, H. Nolla and D. Vaulot (1997) Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991-1994 ENSO event. *Deep Sea Research* **44**:167-192

Campbell, L. and D. Vaulot (1993) Photosynthetic picoplankton community structure in the subtropical north Pacific ocean near Hawaii (station ALOHA). *Deep Sea Research* **40**:2043-2060

Cavender-Bares, K.K. (1999) Size distributions, population dynamics and single cell properties of marine plankton in diverse nutrient environments. Ph. D. Thesis. Civil and Environmental Engineering, MIT, Cambridge

Dusenberry, J.A. and S.L. Frankel (1994) Increasing the sensitivity of a FACScan flow cytometer to study oceanic picoplankton. *Limnology and Oceanography* **39**:206-209

Ferris, M.J. and B. Palenik (1998) Niche adaptation in ocean cyanobacteria. *Nature* **396**:226-228

Field, K.G., D. Gordon, T. Wright, M. Rappè, E. Urbach, K. Vergin and S.J. Giovannoni (1997) Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Applied and Environmental Microbiology* **63**:63-70

Giovannoni, S.J., T.B. Britschgi, C.L. Moyer and K.G. Field (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**:60-63

Goericke, R.E. and D.J. Repeta (1993) Chlorophylls *a* and *b* and divinyl chlorophylls *a* and *b* in the open subtropical North Atlantic Ocean. *Marine Ecology Progress Series* **101**:307-313

Lee, S.-Y., J. Bollinger, D. Bezdicek and A. Ogram (1996) Estimation of the abundance of an uncultured soil bacterial strain by a competitive quantitative PCR method. *Applied and Environmental Microbiology* **62**:3787-3793

Maidak, B.L., et al. (1999) A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Research* **27**:171-173

Mann, E.L. (2000) Trace Metals and the Ecology of Marine Cyanobacteria. Ph.D. Thesis. Biology, MIT and WHOI, Cambridge, MA

Moore, L.R. and S.W. Chisholm (1999) Photophysiology of the Marine Cyanobacterium *Prochlorococcus:* Ecotypic differences among cultured isolates. *Limnology and Oceanography* **44**:628-638

Moore, L.R., G. Rocap and S.W. Chisholm (1998) Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* Ecotypes. *Nature* **393**:464-467

Moyer, C.L., F.C. Dobbs and D.M. Karl (1994) Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Applied and Environmental Microbiology* **60**:871-879

Mutter, G.L. and K.A. Boynton (1995) PCR bias in amplification of androgen receptor alleles, a trinucleotide repeat marker used in clonality studies. *Nucleic Acids Research* **23**:1411-1418

Neveux, J., F. Lantoine, D. Vaulot, D. Marie and J. Blanchot (1999) Phycoerythrins in the southern tropical and equatorial Pacific Ocean: Evidence for new cyanobacterial types. *Journal of Geophysical Research* **104**:3311-3321

Olson, R.J., S.W. Chisholm, E.R. Zettler and E.V. Armbrust (1990a) Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. *Limnology and Oceanography* **35(1)**:45-58

Olson, R.J., E.R. Zettler, M. Altabet, S.W. Chisholm and J. Dusenberry (1990b) Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Research* **37**:1033-1051

Page, R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**:357-358

Palenik, B. (1994) Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Applied and Environmental Microbiology* **60**:3212-3219

Palenik, B. and R. Haselkorn (1992) Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. *Nature* **355**:265-267

Palys, T., L.K. Nakamura and F.M. Cohan (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *International Journal of Systematic Bacteriology* **47**:1145-1156

Polz, M.F. and C.M. Cavanaugh (1998) Bias in template-to-product ratios in multi-template PCR. *Applied and Environmental Microbiology* **64**:3724-3730

Polz, M.F., C. Harbison and C.M. Cavanaugh (1999) Diversity and heterogeneity of epibiotic bacterial communities on the marine nematode *Eubostrichus dianae*. *Applied and Environmental Microbiology* **65**:4271-4275

Rocap, G., L.R. Moore and S.W. Chishom (1999) Molecular Phylogeny of *Prochlorococcus* Ecotypes. In: Charpy, L. and A.W.D. Larkum (ed) Marine Cyanobacteria. Bulletin de l'Institut Océanographique, Monaco, special issue N° 19 (632 pages), p 107-116

Smith, S.W., R. Overbeek, C.R. Woese and W. Gilbert (1994) The genetic data environment: an expandable GUI for multiple sequence analysis. *CABIOS* **10**:671-675

Suzuki, M. and S.J. Giovannoni (1996) Bias caused by templates annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology* **62**:625-630

Swofford, D.L. (1999) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts

Swofford, D.L., G.J. Olsen, P.J. Wadell and D.M. Hillis (1996) Phylogenetic Inference. In: Hillis, D.M., C. Moritz and B.K. Mable (ed) Molecular Systematics. Sinauer, p 407-514

Tay, S.T.-L., H.H. Hemond, L.R. Krumholz, C.M. Cavanaugh and M.F. Polz (2000) Population dynamics of two toluene degrading bacterial species in a contaminated stream assessed by quantitative PCR. submitted

Toledo, G. and B. Palenik (1997) *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. *Applied and Environmental Microbiology* **63**:4298-4303

Toledo, G., B. Palenik and B. Brahamsha (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. *Applied and Environmental Microbiology* **65**:5247-5251

Urbach, E. and S.W. Chisholm (1998) Genetic diversity in *Prochlorococcus* populations flow cytometrically sorted from the Sargasso Sea and the Gulf Stream. *Limnology and Oceanography* **43**:1615-1630

Urbach, E., D. Distel, D.J. Scanlan, J.B. Waterbury and S.W. Chisholm (1998) Rapid diversification of marine picophytoplankton with dissimilar light harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *Journal of Molecular Evolution* **46**:188-201

Urbach, E., D. Robertson and S.W. Chisholm (1992) Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature* **355**:267-270

Veldhuis, M.J.W., G.W. Kraay and W.W.C. Gieskes (1993) Growth and fluorescence characteristics of ultraplankton on a north-south transect in the eastern North Atlantic. *Deep Sea Research II* **40**:609-626

Waterbury, J.B., S.W. Watson, F.W. Valois and D.G. Franks (1986) Biological and ecological characterization of the marine unicellular cyanobacteria *Synechococcus*. *Can. Bull. Fish. Aquat. Sci.* **214**:71-120

West, N.J. and D. Scanlan, J, (1999) Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Applied and Environmental Microbiology* **65**:2585-2591

CHAPTER SIX

Future Directions

A major contribution of this thesis is the identification of sequence similarity clusters of *Prochlorococcus* and *Synechococcus* that may correspond to ecologically distinct populations. However, the ecologically relevant physiological differences between the strain clusters that may allow them to occupy unique niches are not clear in all cases. Further laboratory work, using representative strains from each cluster, will be required to determine the potential roles of nutrients, metals, temperature and other factors in ecotypic differentiation. Several experiments of this sort, including the exploration of potential ecotypic differences in the utilization of organic nutrients, in requirement of cobalt for growth and in the expression of phycoerythrin are currently underway.

Although much discussion has focused on factors that may differentially regulate the growth of the ecotypes, differential mortality may also be an important determinant of distributions. Viruses are known to influence distributions of *Synechococcus*, as natural populations are resistant to their co-occurring phages (Waterbury and Valois 1993). The same is likely true for *Prochlorococcus*, although the search for viruses which infect *Prochlorococcus* is still underway. Selective grazing may also be an important phenomenon in the wild (Christaki, et al. 1999). As little is known about the specific interactions of marine cyanobacteria with their grazers, this area is ripe for exploration.

The ten ecotypic clusters of *Prochlorococcus* and *Synechococcus* identified in chapter 4 may not be an exhaustive set, and the existence additional clusters should be actively explored. Because, as of this writing, only six independent high B/A isolates of

Prochlorococcus exist in culture, continued efforts should be made to obtain new *Prochlorococcus* isolates, with emphasis on high B/A strains. To this end, particular attention should be paid to trying different physical methods of isolation (i.e. flow cytometric sorting, which has been successful for five isolates, see Chapter 2 and Chapter 4, Table1) and alternative media formulations which may select for different strain types. Continued isolation efforts are also warranted for *Synechococcus*, although this study did not exhaustively sample the physiological diversity already present in culture collections. Thus, an immediate next step would entail obtaining ITS sequences from additional *Synechococcus* isolates, in particular high PUB non-motile strains, low PUB motile strains, and psychrophilic strains. It is also of interest to sequence the ITS of some of the strains analyzed using *rpoC1* in order to be able to compare the clusters found with the two loci (Toledo and Palenik 1997, Toledo, et al. 1999).

The phylogenetic affiliation of clone 150m –47 from Chapter 5 remains puzzling. Because the PCR strategy used in generating the clone libraries involved a primer about halfway through the 23S rDNA, (the ITS was sequenced using an internal primer) additional 23S rDNA sequence (at least 1200 bp) could be obtained from this clone. However, this might be of limited utility, as there are not currently a large number of cyanobacterial 23S rDNA sequences publicly available for comparison (i.e. the latest release of the RDP (Maidak, et al. 1999) contains only one, the freshwater *Synechococcus* strain PCC6301). Another strategy might be to examine the ITS sequences of marine *Synechocystis* strains WH8501 and WH8502 (Waterbury and Rippka 1989) to see if they

show a greater sequence similarity to clone 150m-47 than the *Prochlorococcus* and *Synechococcus* sequences do.

Further efforts can also be made towards assessing diversity in natural populations. The data presented in Chapter 5 was the beginning of connecting information about physiological diversity to genetic diversity in field populations of marine cyanobacteria. However, three PCR clone libraries from two water columns from one ocean is probably not a thorough sampling of the existing diversity of marine cyanobacterial populations. If additional libraries are constructed, sampling locations should be chosen with care. The observation that isolates from different oceans have identical ITS sequences suggests that simple geographic location is not the most important component to consider when selecting sampling sites. Rather, additional samples should be from environments with diverse physical conditions, which may select for alternate ecotypes. For example, the Gulf of Aqaba in the Red Sea undergoes winter mixing so deep that *Prochlorococcus* disappear for a period of time most years (Lindell and Post 1995). It would be interesting to examine the genetic makeup of the last *Prochlorococcus* after the onset of mixing, as well as the first populations to come back in the spring. Similarly late summer or early fall in the Sargasso Sea, when waters have been stratified for months, would provide a different water column, and potentially different populations than those presented here (DuRand, et al. 2000).

However, given the drawbacks associated with examining PCR amplified sequences, the next obvious step is the quantitative assessment of the ecotypes. This is possible using signature sequences specific to each ecotype as probes in quantitative or in situ hybridizations (Giovannoni, et al. 1988, DeLong, et al. 1989), or as primers in quantitative PCR (Lee, et al. 1996, Tay, et al. 2000). An examination of a consensus sequences for each cluster over the first ~150 bp of the 23S rDNA (Fig. 1) reveals that such signature sequences can be designed. For example, a 20-mer from positions 18-38 could be used to develop a probe specific for low B/A *Prochlorococcus*, one for high B/A cluster I *Prochlorococcus*, one for high B/A cluster IV *Prochlorococcus*, and one for all four clusters of marine *Synechococcus*. All of these probes have at least one mismatch from other sequences, the low B/A one has two.

It is not possible to find potential probes for all of the combinations of clusters in this region of the 23S rDNA. For example, there is not a sequence that is specific for all of the high B/A *Prochlorococcus* to the exclusion of the low B/A *Prochlorococcus* and *Synechococcus*. It is possible that such a sequence could be found by sequencing more of the 23S rDNA, but that may not be likely given the diverging branching patterns of the high B/A isolates. However, by using a judicious combination of probes it should be possible to quantitate the clusters of interest. Not only will this enable the testing of current hypotheses about the distribution of the high and low B/A ecotypes with depth, but data on the distribution of the ecotypes would undoubtedly generate new hypotheses

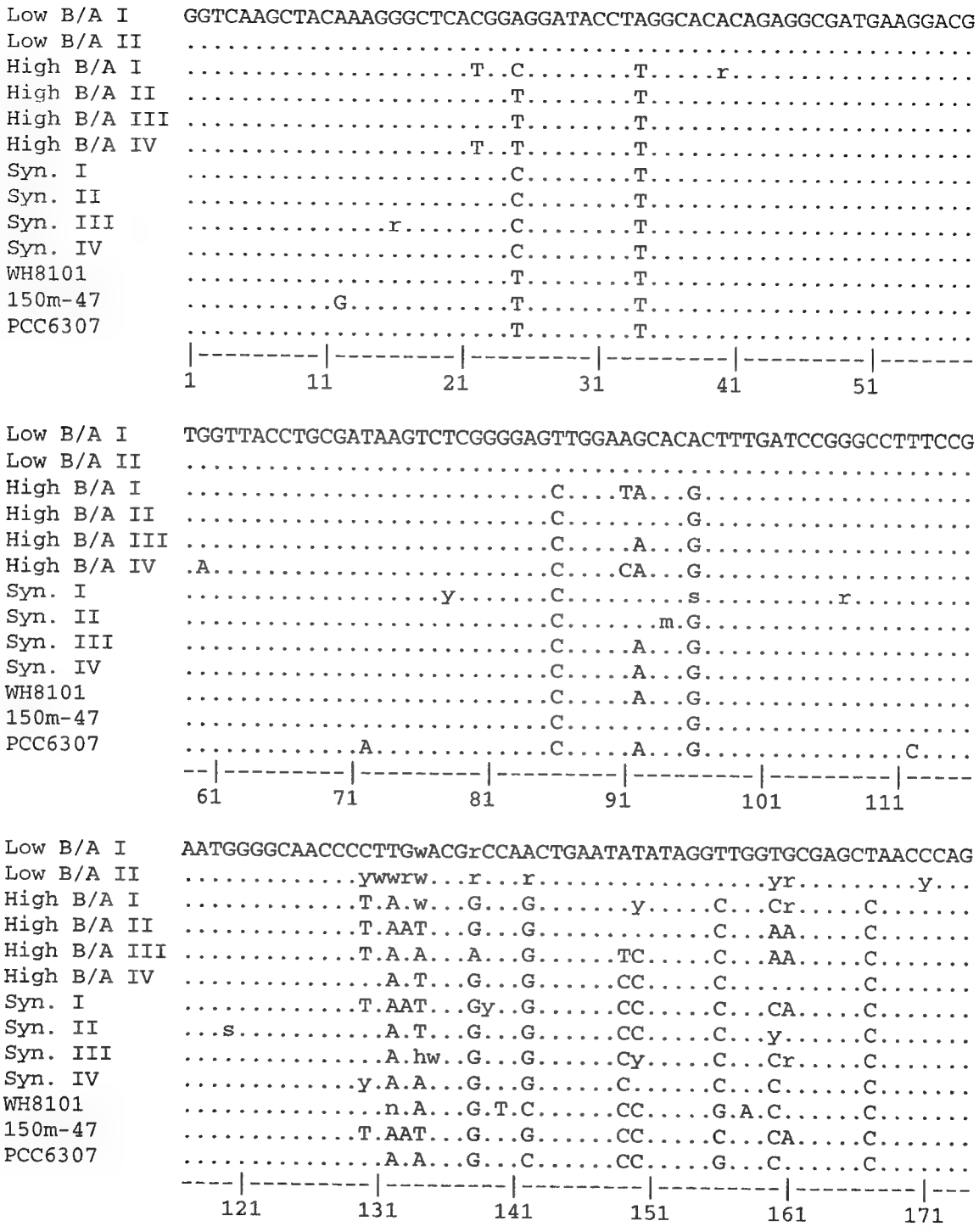


Figure 1. Consensus of partial 23S rDNA sequences of *Prochlorococcus* and *Synechococcus* ecotypes. Sequences from cultures and environmental PCR clone libraries (Chapters 4 and 5) were used to determine a strict consensus sequence for each cluster. Dots at a base position indicate an identical base to that of low B/A cluster I.

that could be explored with laboratory cultures, and might provide hints about the physiology of "types" known only from cloned sequences.

Finally, the complete genome sequence of the low B/A *Prochlorococcus* isolate MED4 will be available by the end of the year (for the status of the project to date see http://bbrp.llnl.gov/jgi/microbial/prochlorococcus_homepage.html). This will provide a wealth of new information and likely drive *Prochlorococcus* research in new directions. Possible benefits from the genome sequence include a better understanding of the photosynthetic apparatus, of nutrient acquisition or metal detoxifying mechanisms, and of the extent of horizontal gene transfer. Further, due to the rapid advances and dropping costs associated with sequencing technology it is likely that complete genome sequences for a high B/A *Prochlorococcus* and a marine *Synechococcus* strain will be determined in the next few years. The comparative genomics approach could be very powerful in understanding the adaptations of each ecotype, and perhaps the costs associated with those adaptations.

REFERENCES

Christaki, U., S. Jacquet, J.R. Dolan, D. Vaulot and F. Rassoulzadegan (1999) Growth and grazing on *Prochlorococcus* and *Synechococcus* by two marine ciliates. *Limnology and Oceanography* **44**:52-61

DeLong, E.F., G.S. Wickham and N.R. Pace (1989) Phylogenetic stains: Ribosomal RNA-based probes for the identification of single cells. *Science* **243**:1360-1363

DuRand, M.D., R.J. Olson and S.W. Chisholm (2000) Phytoplankton population dynamics at the Bermuda Atlantic Time-series Station in the Sargasso Sea. *Deep Sea Research* in press

Giovannoni, S.J., E.F. DeLong, G.J. Olsen and N.R. Pace (1988) Phylogenetic Group-specific oligodeoxynucleotide probes for identification of single Microbial cells. *Journal of Bacteriology* **270**:720-726

Lee, S.-Y., J. Bollinger, D. Bezdicek and A. Ogram (1996) Estimation of the abundance of an uncultured soil bacterial strain by a competitive quantitative PCR method. *Applied and Environmental Microbiology* **62**:3787-3793

Lindell, D. and A.F. Post (1995) Ultraphytoplankton sucession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnology and Oceanography* **40**:1130-1141

Maidak, B.L., et al. (1999) A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Research* **27**:171-173

Tay, S.T.-L., H.H. Hemond, L.R. Krumholz, C.M. Cavanaugh and M.F. Polz (2000) Population dynamics of two toluene degrading bacterial species in a contaminated stream assessed by quantitative PCR. submitted

Toledo, G. and B. Palenik (1997) *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. *Applied and Environmental Microbiology* **63**:4298-4303

Toledo, G., B. Palenik and B. Brahamsha (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. *Applied and Environmental Microbiology* **65**:5247-5251

Waterbury, J.B. and R. Rippka (1989) Subsection I. Order Chroococcales Wettstein 1924, Emend. Rippka et al., 1979. In: Staley, J.T., M.P. Bryant, N. Pfennig and H.J. G. (ed) Bergey's Manual of Systematic Bacteriology v. 3. Williams and Wilkins, Baltimore, p 1728-1745

Waterbury, J.B. and F.W. Valois (1993) Resistance to co-occuring phages enables marine *Synechococcus* communities to to coexist with cyanophages abundant in seawater. *Applied and Environmental Microbiology* **59**:3393-3399

APPENDIX I

Characterization of phycoerythrin genes in the chlorophyll a_2/b_2 containing prokaryote *Prochlorococcus* sp. MIT9303

Claire Ting, Gabrielle Rocap, Jonathan King, and Sallie W. Chisholm

1999 Ting, C., G. Rocap, J. King and S. W. Chisholm. Molecular Characterization of phycoerythrin in the chlorophyll a_2/b_2-containing prokaryote *Prochlorococcus*.sp. MIT9303 pp 225-228 in (Garab, G , Ed) Photosynthesis: mechanisms and effects. Kluwer Academic Publishers, The Netherlands.

CHARACTERIZATION OF PHYCOERYTHRIN GENES IN THE CHLOROPHYLL A_2/B_2-CONTAINING PROKARYOTE, *PROCHLOROCOCCUS* SP. MIT9303

Claire Ting[1], Gabrielle Rocap[1], Jonathan King[1], and Sallie W. Chisholm[1,2]
[1]Dept. of Biology and [2]Dept. of Civil and Environmental Engineering,
Massachusetts Institute of Technology, Cambridge, MA 02139 USA

Key words: aquatic ecosystems, cyanobacteria, phototrophs, phycobiliproteins, sequence analysis, ultrastructure

1. Introduction

Prochlorococcus is responsible for a significant fraction of the photosynthesis in the open ocean. This marine prokaryote often thrives in deep waters where irradiance levels are 0.1% to 1% of surface values. The ability of *Prochlorococcus* to flourish in low light environments may be due to unique functional and structural characteristics of its photosynthetic apparatus. Molecular phylogenies based on 16S rRNA gene analyses indicate that *Prochlorococcus* is a member of the cyanobacterial lineage (1). However, unlike its cyanobacterial counterparts, *Prochlorococcus* has a chlorophyll a_2/b_2 light-harvesting complex (2), and evidence has not been obtained for the presence of functional phycobilisomes. Hess et al. (3) have reported that genes encoding the α- and β-subunits of phycoerythrin (PE) are present in two (SS120, TATL2) of the six *Prochlorococcus* isolates which they examined. Both SS120 (120 m, Sargasso Sea) and TATL2 (30 m, Tropical Atlantic) were isolated from the Atlantic. Three other isolates were also from the Atlantic (10-30 m), and one was from the Mediterranean (5 m) (3). Although the PE genes are expressed at low levels in SS120 and the gene products appear to bind chromophores (3), the function of PE in this organism remains unknown. Furthermore, it has not been determined if the unique characteristics of the SS120 PE sequence (3) are conserved between *Prochlorococcus* isolates.

We have in culture another *Prochlorococcus* isolate, MIT9303, from 100 m in the Sargasso Sea. Among all of the cultured isolates, it has the highest degree of 16S rRNA sequence similarity to marine *Synechococcus*, the group from which *Prochlorococcus* may have evolved (1,4). Using Southern blot analyses and/or PCR amplification, we detected the α- and β-PE genes in MIT9303, but not in two other isolates from different sites. The predicted amino acid sequences of PE of MIT9303 and SS120 (3) differ in their number of putative chromophore binding sites and in many other positions. Our ultrastructural studies did not reveal phycobilisome-like structures associated with the internal membranes of MIT9303. At present, the physiological role of PE in *Prochlorococcus* remains to be established.

2. Procedure

Prochlorococcus cells were grown in Sargasso Sea water enriched with nutrients, at 21°C on a 14 hr light (10 µmol photons m^{-2} s^{-1}) / 10 hr dark cycle. Cells were prepared for electron microscopy as described in (5). For our Southern blots, the SS120 PE genes were amplified, labeled with fluorescein (ECL Random Prime Labeling and Detection Systems, Amersham, IL), and utilized as the probe. Genomic DNA was purified (6), digested with *HindIII*, separated on an agarose gel, and transferred to nitrocellulose. These membranes were hybridized with the PE gene probe, blocked, incubated with anti-fluorescein-HRP conjugate and detection solutions, and exposed to autoradiography film. For the characterization of PE genes, genomic DNA was isolated from MIT9303, and genes encoding the β- and α-subunits of PE were amplified using primers specific for *cpeB* (forward primer: ATGCTTGATGCATTCTCAAG) and *cpeA* (reverse primer: AAGGCATTAATAAGGTAATC) of SS120 (3). PCR products were separated on an agarose gel, purified, and cloned using the pCR-Script kit (Stratagene, CA). Plasmid DNA from several independent clones was sequenced on an automated sequencer (LICOR, NE). Sequences were aligned using the Genetic Data Environment.

3. Results and Discussion

Genes encoding PE were detected previously in two out of six *Prochlorococcus* isolates examined (3), and Hess et al. (1996) suggested that during evolution, the development of a chlorophyll *a/b* light-harvesting complex in *Prochlorococcus* led to the disappearance of phycobilisomes in some lines (3). We used Southern blot analyses to confirm that *cpeA* and *cpeB* are absent from MED4, as previously reported (3), and are also absent from NATL2A and MIT9313 (data not shown). These latter isolates were originally obtained from the Atlantic, with NATL2A originating from 10 m (North Atlantic) and MIT9313 from 135 m (Gulf Stream). However, *cpeA* and *cpeB* are present in MIT9303. Using primers specific for the SS120 PE genes and the PCR technique, we amplified a region of about 1000 base pairs from the genomic DNA of MIT9303. Sequencing confirmed the presence of *cpeA* and *cpeB* on this fragment.

The MIT9303 *cpeA* gene is similar to *cpeA* of SS120 (3), and does not contain additional codons or deletions. The *cpeA* genes of both MIT9303 and SS120 are slightly smaller than *cpeA* of marine *Synechococcus* sp. strain WH7803 (7) because of two stretches of deletions (12-15 nucleotides each) in less-conserved regions of the gene. Although the *cpeB* gene of MIT9303 is also similar to *cpeB* of SS120 (3), it is missing a codon which encodes Asn-156 in SS120. In addition, both the MIT9303 and SS120 *cpeB* genes are missing two codons which encode Lys-157 and Met-158 in WH7803 (7).

The predicted amino acid sequences of the β-subunits of MIT9303 (sequence not shown) and SS120 (3) share approximately 73% amino acid identity. In the β-subunit, the four cysteines (Cys-50, -61, -82, -165), which are important in chromophore binding (7-10), are conserved in MIT9303. In addition, residues which are involved in chromophore interaction (Arg-77,-78, Ala-81, Arg-84, Asp-85) (9) are also conserved.

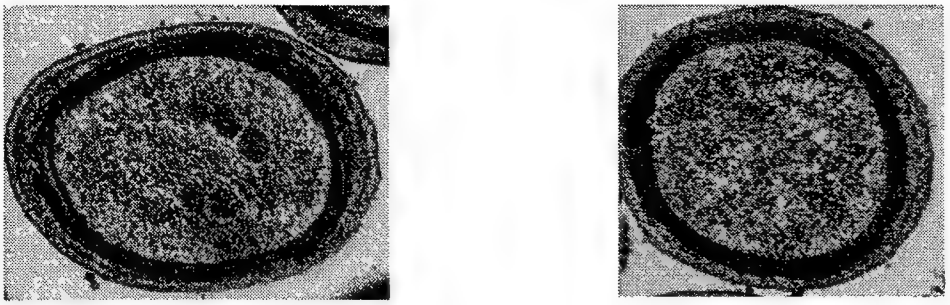
The α-PE subunits of MIT9303 and SS120 share approximately 65% amino acid identity (Fig. 1). An important difference in the predicted amino acid sequence of α-PE of MIT9303 involves the presence of two cysteine residues (Cys-73, Cys-130), which are putative chromophore binding sites (Fig. 1). The α-PE of SS120 contains only Cys- 73, and has an Asp residue substituted for Cys-130 (3). While Cys-73 is homologous to a bilin attachment site (Cys-88) that is conserved in almost all phycobiliproteins (9), Cys-130 is homologous to a chromophore attachment site at Cys-140 in organisms including *Synechococcus* sp. strains WH7803 (Class I PE) (7), and WH8103 and WH8020 (Class II PE) (8). The PE (α+β subunits) of MIT9303 has a total of five putative chromophore attachment sites, a characteristic associated with Class I PE (7-10). In the α-PE of both MIT9303 and SS120 (3), Arg-75 and Asp-76, which are highly conserved residues involved in maintaining proper chromophore conformation (9), are also present.

```
MIT9303   M K S A V T T V I T A A D A A G R F P D I S D L K A
SS120     M K S T V T T V I A S A D A A G R F P T I S D I E S
          27
MIT9303   V K A S F D R A A A R M E A A E K L A S G I D N V T
SS120     V K G S F D R A K D R L E A A E K L S I H I D R F T
          53                                  *   • •
MIT9303   A D A L K A V Y S D G K Y D L A T R D K C A R D I N
SS120     S Q A L D Y V Y G T E N Y E Q A N K D K C S R D I H
          79
MIT9303   H Y L R L I N Y C L I A G S T G P L D E W G I A G V
SS120     H Y L R L I N Y C L V T G G T G P L D E W G I A G M
          105                                         *
MIT9303   R E V F R T L G I P T S A Y I E A F S Y I R E R V C
SS120     R E V I R I Q L L P T A A Y I E A F T Y I R D N L D
          131
MIT9303   V P R D M D Q Q A A N E F K D L L
SS120     I P N D M G Q Q A G A E F K N L L
```

Figure 1. Predicted amino acid sequences of α-PE of MIT9303 and SS120 (3). The last eight amino acid residues of the MIT9303 α-PE sequence have not been shown because their corresponding nucleotides were used as the primer in the original amplification of the PE gene in this isolate. (*) indicate conserved Cys that are putative chromophore attachment sites. (•) represent functionally important residues discussed in the text.

In order to determine whether the intracytoplasmic membranes of MIT9303 are associated with phycobilisome-like structures, we examined the ultrastructure of this isolate. In MIT9303, the intracytoplasmic lamellae are organized into parallel bands, which extend the length of the cell and/or form concentric circles near the periphery of the cytoplasmic membrane (Figs. 2A, 2B). The lamellae in these preparations were tightly appressed and lacked phycobilisome-like structures (Figs. 2A, 2B). This organization and the overall ultrastructure of MIT9303, are similar to what we have observed for other *Prochlorococcus* isolates (5,11).

2A. 2B.



Figures 2A and 2B. Transmission electron micrographs of thin sections of MIT9303. A.) Longitudinal section B.) Cross-section

The possession of *cpeA* and *cpeB* is a characteristic that is not conserved among *Prochlorococcus* isolates. An understanding of the driving forces which may have resulted in the acquisition and/or loss of these genes is linked intimately with further studies on the function of PE in *Prochlorococcus*. These studies will also allow us to determine whether the presence of *cpeA* and *cpeB* in MIT9303 and SS120 confers a biological advantage under specific light and/or nutrient conditions.

4. Addendum

This work was supported by a NSF Postdoctoral Research Fellowship in Biosciences Related to the Environment awarded to C. Ting, and by a NIH GM 17,980 to J. King. G. Rocap and S.W. Chisholm would like to acknowledge the support of a NASA grant.

References

1. Urbach E, Robertson D, Chisholm SW (1992) Nature 355:267-270
2. Partensky F, LaRoche J, Wyman K, Falkowski PG (1997) Photosyn. Res. 51:209-222
3. Hess WR, Partensky F, van der Staay GWM, Garcia-Fernandez JM, Borner T, Vaulot D (1996) PNAS 93:11126-11130
4. Urbach E, Scanlan DJ, Distel DL, Waterbury JB, Chisholm SW (1998) J. Mol. Evol. 46:188-201
5. Ting C, King J, Chisholm SW (1998) Manuscript in preparation
6. Ausubel F, Brent R, Kingston RE, Moore D, Seidman JG, Smith JA, Struhl K (1995) Short Protocols in Molecular Biology, Third Edition, John Wiley and Sons, USA
7. Newman J, Mann N, Carr NG (1994) Plant Mol. Biol. 24:679-683
8. de Lorimier RM, Chen CCJ, Glazer AN (1992) Plant Mol. Biol. 20:353-356
9. Apt KE, Collier JL, Grossman AR (1995) J. Mol. Biol. 248:79-96
10. Ong LJ, Glazer AN (1991) J. Biol. Chem. 266:9515-9527
11. Chisholm SW, Olson RJ, Zettler ER, Goericke R, Waterbury J, Welschmeyer N (1988) Nature 334:340-343

APPENDIX II

Phylogeny of *Prochlorococcus* using 16S rRNA revisited

Since the publication of Chapters 2 and 3 of this thesis (Moore, et al. 1998, Rocap, et al. 1999), additional marine cyanobacterial 16S rRNA genes have been sequenced, both from cultures (D. Distel and J. Waterbury, unpublished data) and from the environment (West and Scanlan 1999). Further, sequences from additional freshwater *Synechococcus* suggest that the *Synechococcus* group as a whole is paraphyletic within the cyanobacterial radiation, and that the strain *Synechococcus* PCC6301, which was used as the outgroup in Chapters 2 and 3, is in a different lineage than the marine picoplankton clade (Honda, et al. 1999).

Here I present a phylogenetic analysis of the *Prochlorococcus* 16S rDNA sequences from Chapters 2 and 3 with the new sequences now available. The trees are rooted with the freshwater cyanobacterial strain PCC6307 since it is more closely related to the marine picoplankton clade than *Synechococcus* PCC6301 (Honda, et al. 1999). I also present an extended version of the methods used to acquire the sequences in Chapters 2 and 3, which were not elaborated in those chapters, because of the space constraints imposed by the journal in Chapter 2 and because Chapter 3 was written to review the state of knowledge about the ecotypes of *Prochlorococcus* at the end of 1998.

METHODS

Culturing and DNA isolation. Cultures were grown in 60 ml acid washed polycarbonate bottles in Pro2 media (see Chapter 4). Genomic DNA was isolated from late exponential phase cultures using a protocol adapted from standard methods (Ausubel,

et al. 1992). Cells were spun in an ultracentrifuge for 20 min at 10,000 rpm at 4° C. The pellet was transferred to a 1.5 ml tube and spun in a microcentrifuge for 10 min at max speed. The supernatant was removed, the pellet was resuspended in 430 µl TE and 17 µl 50 mg/ml lysozyme (final concentration 2 mg/ml) was added. The sample was incubated at 37° C for 30 min and 30 µl 10% SDS, 60 µl 5% Sarcosyl and 60 µl 5% TritonX-100 were added (all final concentration 0.5%). The sample was incubated at 37° C for 1 hour and 100 µl 5M NaCl and 80 µl CTAB/NaCl (10% CTAB, 0.7 M NaCl) were added and the sample was incubated in a 65° C water bath for 10 min. The sample was extracted with an equal volume of 24:1 chloroform:isoamyl alcohol, then with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol, and then repeatedly (usually twice more) with 24:1 chloroform:isoamyl alcohol until no interface was visible between the aqueous and the organic phases. DNA was precipitated with 0.6 volumes isopropyl alcohol and stored at 4° or -20° overnight. The next day the sample was spun for 10min at 4° C in a microfuge at max speed, the pellet was washed twice in 70% ethanol and dried on the bench or in a speed vacuum (using no heat setting), and resuspended in 30-100 µl TE. For all samples a no-cell blank was carried through the extraction procedure in parallel (starting with 430 µl TE) and used as a control in PCR reactions.

Amplification and cloning of 16S rDNA. The 16S rDNA was amplified using universal primers 8-27f (AGAGTTTGATCCTGGCTCAG) and 1504-1486r (CTTGTTACGACTTCACCCC). Reactions were done in a 25 µl volume with final

concentrations of reactants as follows: 0.25 mM dNTPs, 0.5 mM each primer, 01.-1 μg template DNA, 1X Pfu Buffer and 0.1-0.5 U of the high fidelity polymerase Pfu (Stratagene). Cycling parameters were 94° for 5 min followed by 30 cycles of (92° 1min, 60° 1min, 72° 2 min) and a final extension at 72° for 10 min in a PTC-100 PCR machine (MJ Research). PCR reactions were performed in quintuplicate and pooled and purified using QIA quick kits (Qiagen).

PCR products were cloned using pCR-Script kits (Stratagene). Plasmids were prepared from clones using a standard alkaline lysis minprep method (Ausubel, et al. 1992) and screened by digestion with Pvu II, which cuts the vector pCR-Script twice, once on either side of the insertion site, and does not cut *Prochlorococcus* 16S rDNA sequences. Plasmids with appropriately sized inserts were prepared for sequencing using a high purity miniprep involving a phenol/chloroform extraction and a PEG/NaCl precipitation.

Sequencing. All sequencing reactions were performed using Epicentre Technologies Long-Read cycling sequencing kits (Sequi-Therm, Sequi-Therm Excel, or Sequi-Therm Excel II) according to the manufacturer's instructions using dye labeled primers, and run on a LI-COR Model 4000 automated sequencer (LI-COR, Lincoln, NE).

For the sequences in Chapter 2 (MIT9302, MIT9303, MIT9312, MIT9313), at least eight clones from each culture were sequenced independently on one strand using primers

M13F and M13R and plasmid DNA as the template. For second strand sequencing plasmids were pooled in equimolar amounts and sequenced using internal primers.

For the sequences in Chapter 3 (MIT9201, MIT9202, MIT9211, MIT9215), at least three clones per culture were sequenced independently on one strand using primers M13F and M13R and plasmid DNA as the template. For second strand sequencing the 16S rDNA was amplified from genomic DNA as described above and sequenced using internal primers.

Phylogenetic analyses. Sequences were manually edited and aligned using the Genetic Data Environment (Smith, et al. 1994). Phylogenetic analyses were performed with PAUP* v. 4b2a (Swofford 1999) and employed 1076 unambiguously aligned and determined positions of the 16S rDNA. Distance trees were inferred using the HKY85 model of nucleotide substitution. Heuristic searches and bootstrap analyses (100 resamplings) utilized random addition and tree-bisection reconnection (TBR) branch swapping methods. Phylogenetic trees were visualized with TREEVIEW (Page 1996).

Nucleotide sequence accession numbers. Sequences from Chapter 3 have been deposited in GenBank with the following accession numbers; MIT9201: AF115268 MIT9202: AF115269 MIT9211: AF115270 and MIT9215: AF115271.

RESULTS

Phylogenetic analyses employing 1076 bp of the 16S rDNA and including sequences from chapter 2 and 3 as well as sequences from marine *Synechococcus* WH8012 and WH8112 (unpublished data courtesy of D. Distel and J. Waterbury) and environmental sequences from the North Atlantic (West and Scanlan 1999) demonstrate many of the features identified in chapter 4 using the more variable ITS. The division of the low B/A *Prochlorococcus* clade into two clusters is well supported, as are the high B/A clusters I and IV (Fig. 1). The motile *Synechococcus* isolate WH8112 is closely related to motile isolate WH8103, consistent with results from the ITS. In contrast *Synechococcus* WH8012 is not specifically associated with the other *Synechococcus* in this analysis, consistent with results from the ITS placing WH8012 in a separate cluster from WH7805 and WH8103 (Chapter 4).

As described previously (West and Scanlan 1999), the environmental sequences from the North Atlantic fall into two groups, one of which is closely related to MED4 (low B/A cluster I) and the other is closely related to NATL2A (high B/A cluster I) (Fig. 1). Environmental sequences SAR 100 and SAR139 group with the motile *Synechococcus* isolates in cluster II. The position of environmental sequence SAR7 remains uncertain, as bootstrap analyses do not provide consistent support for its placement in either the *Prochlorococcus* or the *Synechococcus* clade.

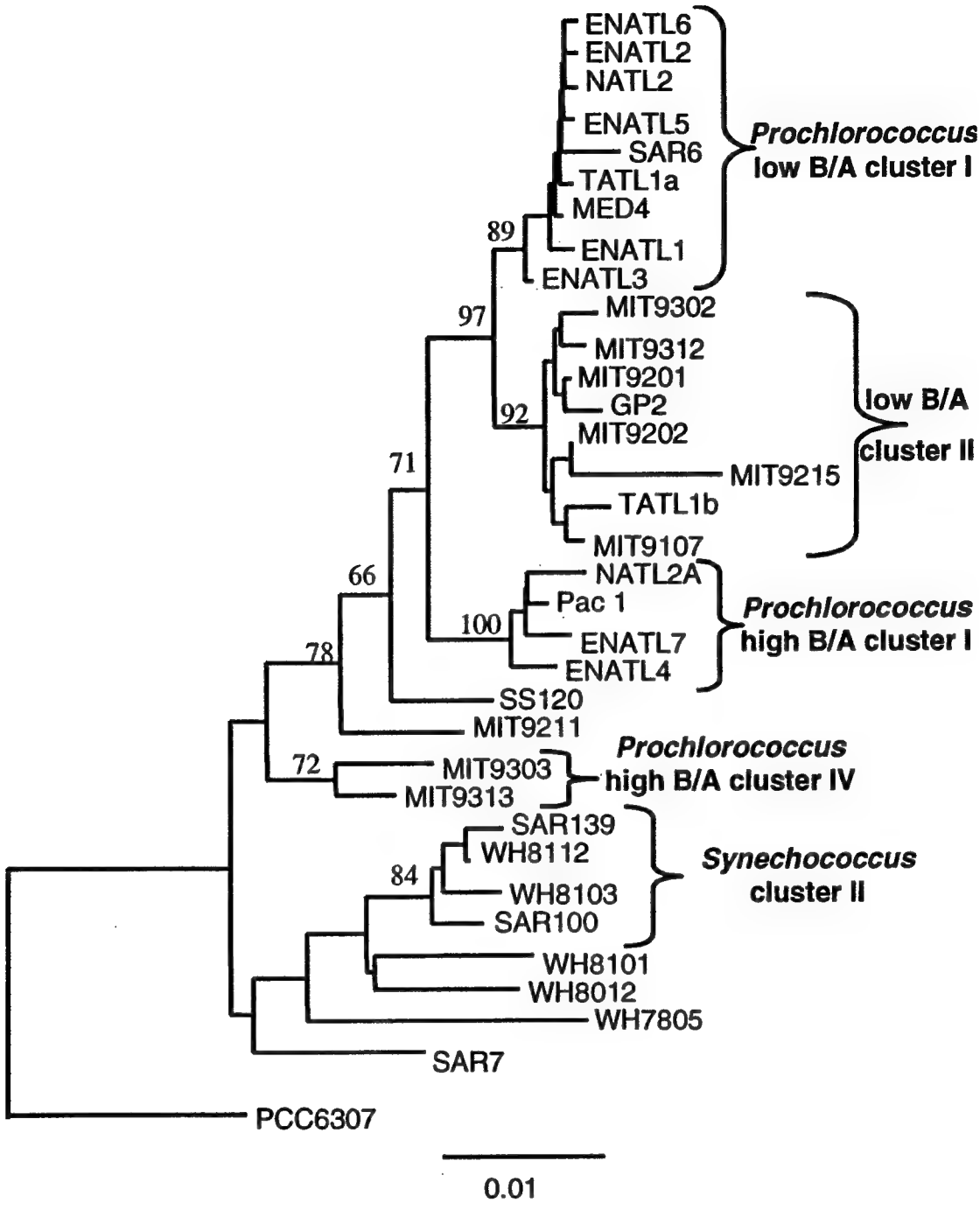


Figure 1. Evolutionary relationships of *Prochlorococcus* and *Synechococcus* 16S rDNA sequences and environmental sequences from the Sargasso Sea (SAR) and the North Atlantic (ENATL). Phylogenetic framework inferred using a distance criterion using minimum evolution as the objective function. Bootstrap values (from 100 resamplings) are presented to the left of each node, values below 50 are not shown. The tree is rooted with Cyanobium cluster *Synechococcus* PCC6307.

REFERENCES

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl (1992) Short Protocols in Molecular Biology. John Wiley & Sons, New York

Honda, D., A. Yokota and J. Sugiyama (1999) Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *Journal of Molecular Evolution* **48**:723-739

Moore, L.R., G. Rocap and S.W. Chisholm (1998) Physiology and Molecular Phylogeny of Coexisting *Prochlorococcus* Ecotypes. *Nature* **393**:464-467

Page, R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**:357-358

Rocap, G., L.R. Moore and S.W. Chishom (1999) Molecular Phylogeny of *Prochlorococcus* Ecotypes. In: Charpy, L. and A.W.D. Larkum (ed) Marine Cyanobacteria. Bulletin de l'Institut Océanographique, Monaco, special issue N° 19 (632 pages), p 107-116

Smith, S.W., R. Overbeek, C.R. Woese and W. Gilbert (1994) The genetic data environment: an expandable GUI for multiple sequence analysis. *CABIOS* **10**:671-675

Swofford, D.L. (1999) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts

West, N.J. and D. Scanlan, J, (1999) Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Applied and Environmental Microbiology* **65**:2585-2591

Appendix III

Secondary structures of sequences in chapters 2

One of the features of ribosomal RNA that makes it amenable to phylogenetic analysis is its conserved secondary structure which permits alignment of sequences across large evolutionary distances (Olsen and Woese 1993). In order to check that base pair changes between strains of *Prochlorococcus* would retain a conserved secondary structure (Gutell, et al. 1994), and to have a better understanding of potential primer and probe design sites, putative secondary structures were plotted for the 16S rDNA sequences obtained in chapter 2 (Fig. 1, A, B, C, D). The overall structure was derived from that for *Synechococcus* PCC6301 (Gutell 1994) and the *Prochlorococcus* sequences were plotted on this framework using the program CARD (Winnepenninckx, et al. 1995).

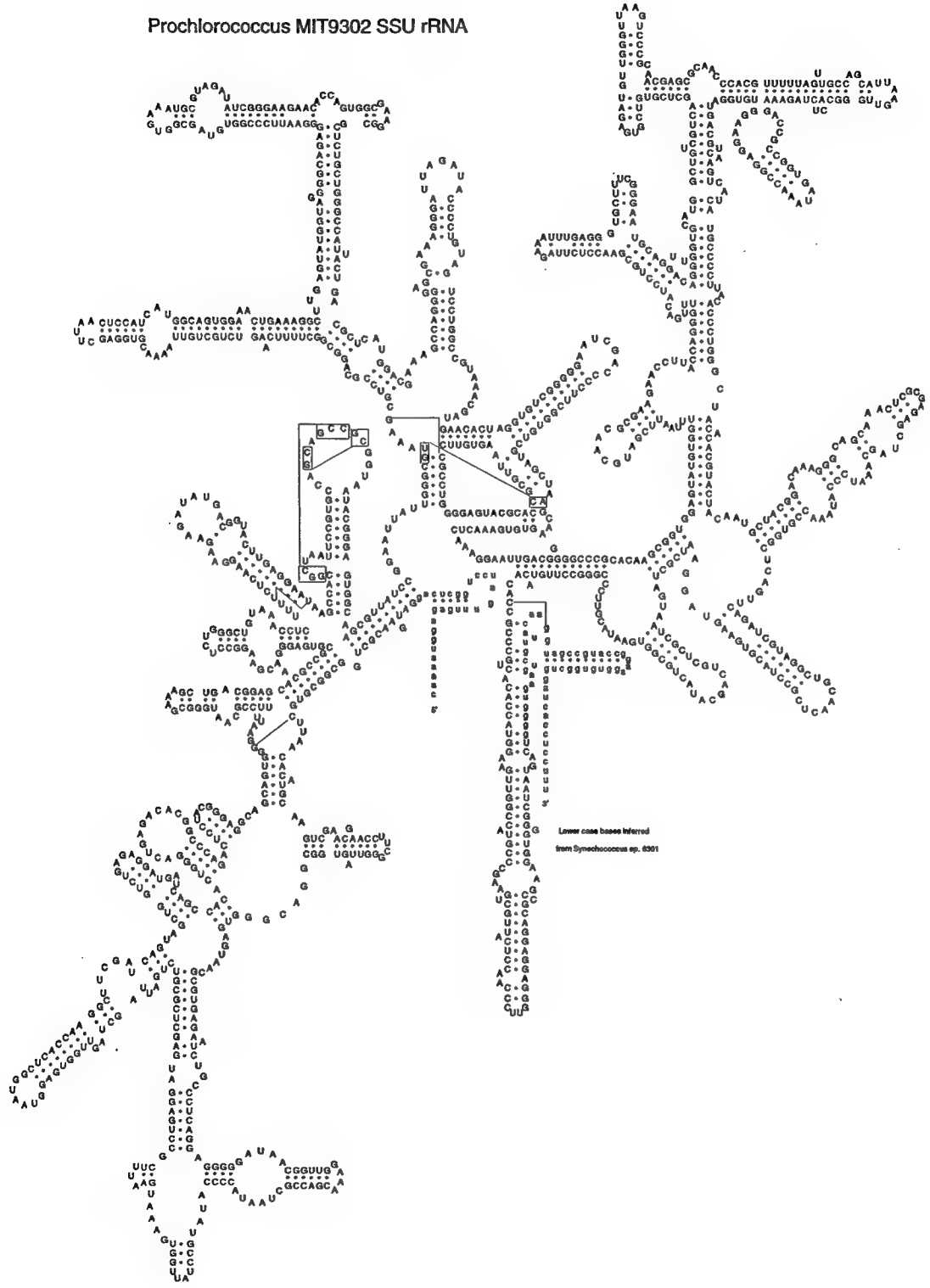
REFERENCES

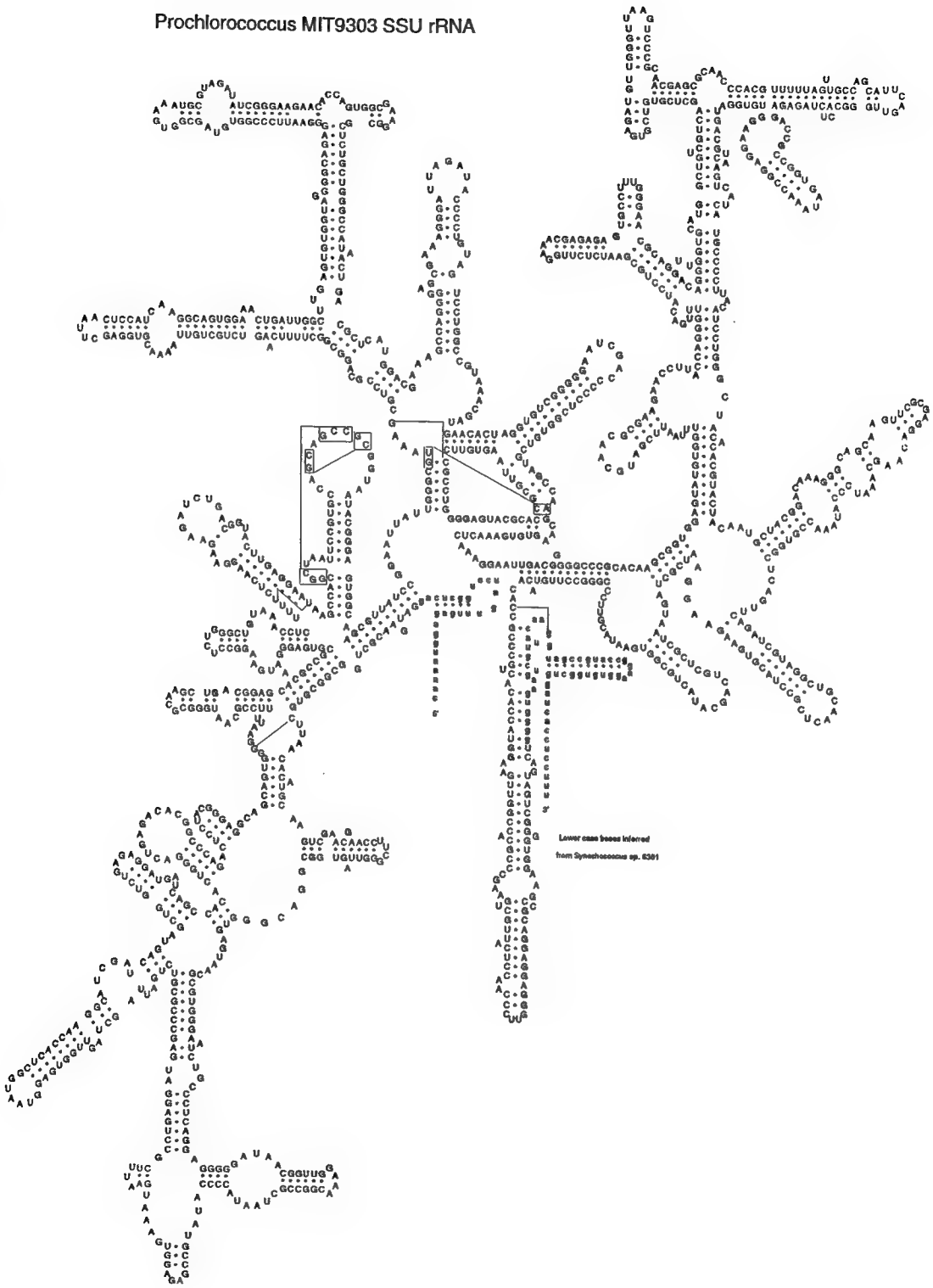
Gutell, R.R. (1994) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. *Nucleic Acids Research* **22**:3502-3507

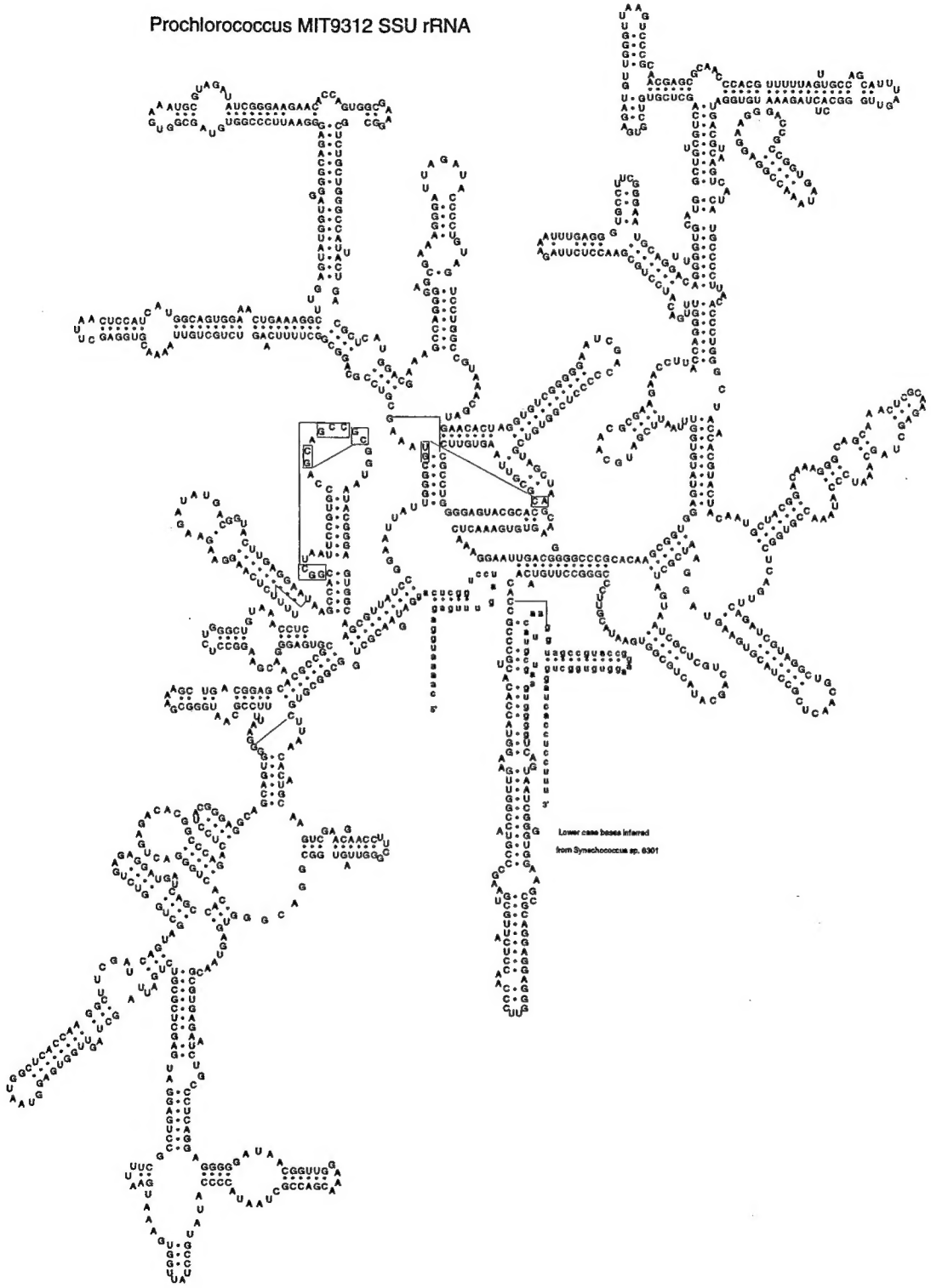
Gutell, R.R., N. Larsen and C.R. Woese (1994) Lessons from an Evolving rRNA: 16S and 23S rRNA Structures from a Comparative Perspective. *Microbiological Reviews* **58**:10-26

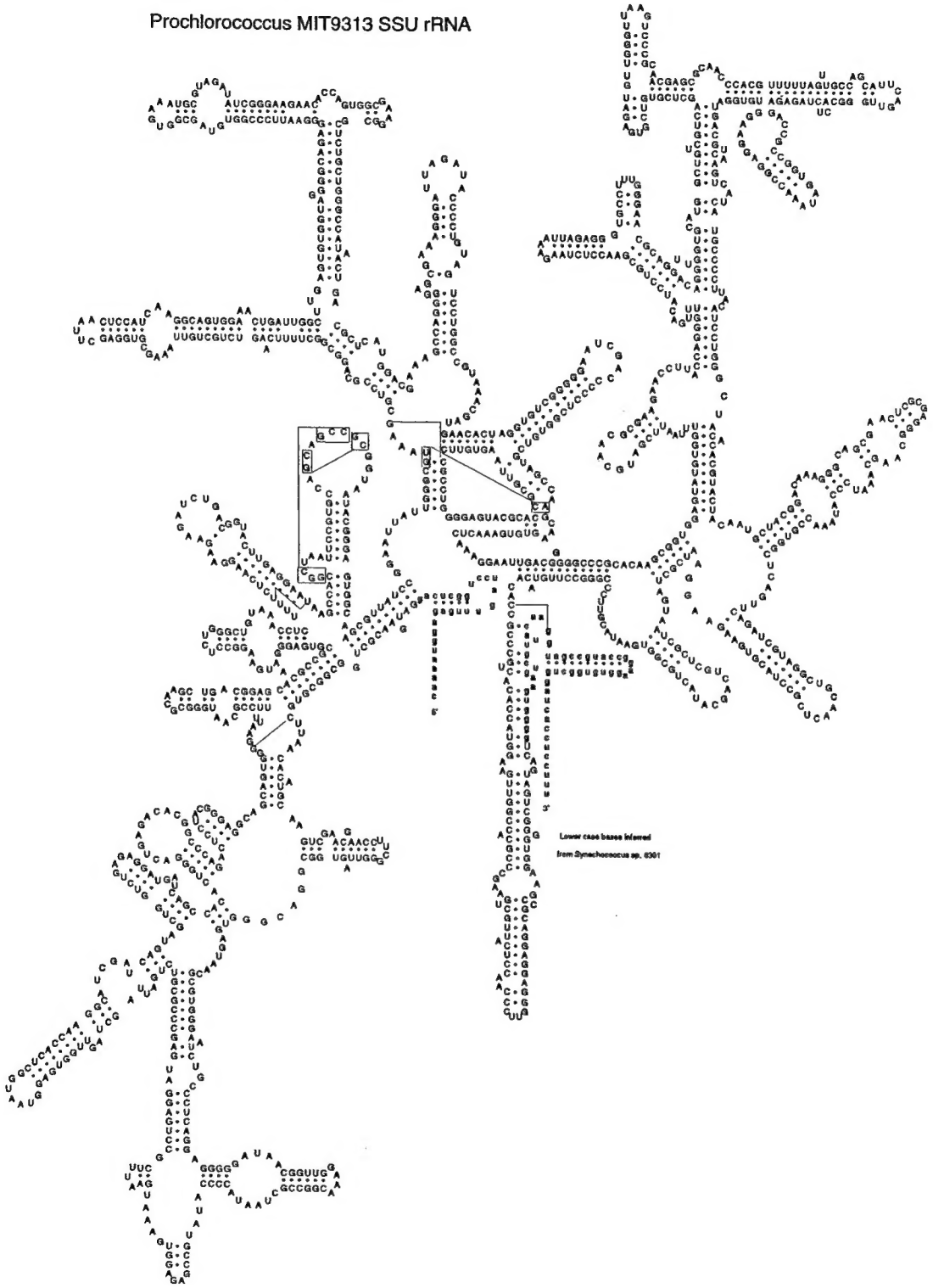
Olsen, G.J. and C.R. Woese (1993) Ribosomal RNA: a key to phylogeny. *FASEB Journal* **7**:113-123

Winnepenninckx, B., Y. Van de Peer, T. Backeljau and R. De Wachter (1995) CARD: a drawing tool for RNA secondary structure models. *Biotechniques* **18**:1060-1063


Prochlorococcus MIT9302 SSU rRNA
Lower case bases inferred
from Synechococcus sp. 6301


Prochlorococcus MIT9303 SSU rRNA
Lower case bases inferred
from Synechococcus sp. 6301


Prochlorococcus MIT9312 SSU rRNA
Lower case bases inferred
from Synechococcus sp. 6301


Prochlorococcus MIT9313 SSU rRNA
Lower case bases inferred
from Synechococcus sp. 6301

Document Library

Distribution List for Technical Report Exchange—November 1999

University of California, San Diego
SIO Library 0175C
9500 Gilman Drive
La Jolla, CA 92093-0175

Hancock Library of Biology & Oceanography
Alan Hancock Laboratory
University of Southern California
University Park
Los Angeles, CA 90089-0371

Gifts & Exchanges
Library
Bedford Institute of Oceanography
P.O. Box 1006
Dartmouth, NS B2Y 4 A2
CANADA

NOAA/EDIS Miami Library Center
4301 Rickenbacker Causeway
Miami, FL 33149

Research Library
U.S. Army Corps of Engineers
Waterways Experiment Station
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Institute of Geophysics
University of Hawaii
Library Room 252
2525 Correa Road
Honolulu, HI 96822

Marine Resources Information Center
Building E38-320
MIT
Cambridge, MA 02139

Library
Lamont-Doherty Geological Observatory
Columbia University
Palisades, NY 10964

Library
Serials Department
Oregon State University
Corvallis, OR 97331

Pell Marine Science Library
University of Rhode Island
Narragansett Bay Campus
Narragansett, RI 02882

Working Collection
Texas A&M University
Dept. of Oceanography
College Station, TX 77843

Fisheries-Oceanography Library
151 Oceanography Teaching Bldg.
University of Washington
Seattle, WA 98195

Library
R.S.M.A.S.
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Maury Oceanographic Library
Naval Oceanographic Office
Building 1003 South
1002 Balch Blvd.
Stennis Space Center, MS 39522-5001

Library
Institute of Ocean Sciences
P.O. Box 6000
Sidney, B.C. V8L 4B2
CANADA

National Oceanographic Library
Southampton Oceanography Centre
European Way
Southampton SO14 3ZH
UK

The Librarian
CSIRO Marine Laboratories
G.P.O. Box 1538
Hobart, Tasmania
AUSTRALIA 7001

Library
Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead
Merseyside L43 7 RA
UK

IFREMER
Centre de Brest
Service Documentation–Publications
BP 70 29280 PLOUZANE
FRANCE

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO. MIT/WHOI 00-23	2.	3. Recipient's Accession No.

4. Title and Subtitle
Genetic Diversity and Ecotypic Differentiation in the Marine Cyanobacteria *Prochlorococcus* and *Synechococcus*

5. Report Date February 2000

6.

7. Author(s) Gabrielle Rocap

8. Performing Organization Rept. No.

9. Performing Organization Name and Address
MIT/WHOI Joint Program in Oceanography/Applied Ocean Science & Engineering

10. Project/Task/Work Unit No. MIT/WHOI 00-23

11. Contract(C) or Grant(G) No.
(C) NAG5-3727
(G) OCE-9820035

12. Sponsoring Organization Name and Address
National Science Foundation
National Aeronautics and Space Administration

13. Type of Report & Period Covered Ph.D. Thesis

14.

15. Supplementary Notes
This thesis should be cited as: Gabrielle Rocap, 2000. Genetic Diversity and Ecotypic Differentiation in the Marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Ph.D. Thesis. MIT/WHOI, 00-23.

16. Abstract (Limit: 200 words)

Primary production in the open ocean is dominated by the closely related cyanobacterial genera *Prochlorococcus* and *Synechococcus*. This thesis explores the relationship between physiological and genetic diversity in cultured isolates of *Prochlorococcus* and *Synechococcus* and examines genetic diversity in natural cyanobacterial populations to better understand the ecology of this globally important clade.

The analysis of 16S ribosomal RNA gene (rDNA) sequences of a spectrum of *Prochlorococcus* isolates from the world's oceans suggested that *Prochlorococcus* could be divided into two distinct ecotypes, designated high and low B/A because of their differing chl b/a_2 ratios. To further resolve the ecotypes, the internal transcribed spacer region (ITS) between the 16 and 23S rRNAs was sequenced in isolates of *Prochlorococcus* and *Synechococcus*. Phylogenetic analysis identified ten sequence clusters, four of *Synechococcus* and six of *Prochlorococcus*, each of which likely corresponds to an ecologically distinct population.

Genetic diversity of natural populations of *Prochlorococcus* and *Synechococcus* was investigated in summer and winter waters of the Sargasso Sea by constructing clone libraries from PCR amplified ITS sequences, which were screened by restriction fragment length polymorphism (RFLP). Based on their phylogenetic relationships the sequences were assigned to eight clusters, seven of which had been previously identified.

17. Document Analysis a. Descriptors
Cyanobacteria
Prochlorococcus
Evolution

b. Identifiers/Open-Ended Terms

c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Approved for publication; distribution unlimited.	UNCLASSIFIED	152
	20. Security Class (This Page)	22. Price

(See ANSI-Z39.18) *See Instructions on Reverse* OPTIONAL FORM 272 (4-77)
(Formerly NTIS-35)
Department of Commerce